AD-A077 065 MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTE--ETC F/G 5/7
TOWARD A COMPUTATIONAL THEORY OF INDIRECT SPEECH ACTS.(U)
OCT 79 G P BROWN N00014-75-C-0661
UNCLASSIFIED MIT/LCS/TR-223 NL

1 OF 3
AD A 077065



LEVEL II

LABORATORY FOR
COMPUTER SCIENCE



MASSACHUSETTS
INSTITUTE OF
TECHNOLOGY

AD A077065

MIT/LCS/TR-223

12

TOWARD A COMPUTATIONAL THEORY OF INDIRECT SPEECH ACTS

DDC
RECEIVED
NOV 19 1979
E

Gretchen P. Brown

This document has been approved for public release and sale; its distribution is unlimited.

This research was supported by
the Advanced Research Projects Agency of
the Department of Defense and was monitored by
the Office of Naval Research
under Contract No. N00014-75-C-0661

DDC FILE COPY

545 TECHNOLOGY SQUARE, CAMBRIDGE, MASSACHUSETTS 02139

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (14) MIT/LCS/TR-223	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) (6) Toward a Computational Theory of Indirect Speech Acts	5. TYPE OF REPORT & PERIOD COVERED (9) Interim rept.
6. PERFORMING ORG. REPORT NUMBER MIT/LCS/TR-223	
7. AUTHOR(s) (10) Gretchen P. Brown	8. CONTRACT OR GRANT NUMBER(s) (15) N00014-75-C-0661
9. PERFORMING ORGANIZATION NAME AND ADDRESS MIT/Laboratory for Computer Science 545 Tecnology Square Cambridge, MA 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS ARPA/Department of Defense 1400 Wilson Boulevard Arlington, VA 22209 (11)	12. REPORT DATE October 1979
	13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) ONR/Department of the Navy Information Systems Program Arlington, VA 22217 (12) 196	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is 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)

speech acts
indirect speech acts
conversational maxims
natural language
computational linguistics
semantics
English dialogue
pragmatics
language recognition

409 648

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

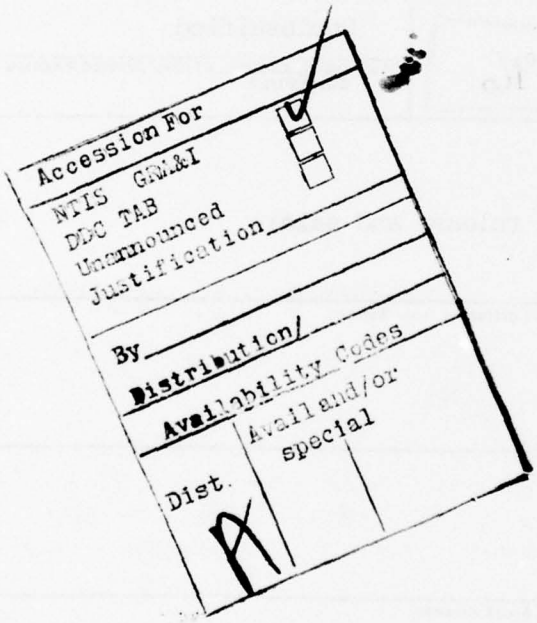
The variety of surface forms that may be used to convey a given speech act pose a major problem in modelling task-oriented (and other) dialogues. Many such forms are so-called indirect speech acts, that is, surface form does not correspond to the (or one) intended speech acts. While this topic has received extensive attention from linguistics, their concerns have not usually been computationally motivated. In this paper, I present a non-computational analysis of indirect speech act forms with an eye to computational considerations.

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

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

79 11 16 049

20. The paper is divided into two parts. Part 1 presents categories and rules for indirect speech acts, justified where possible by traditional linguistic arguments. The second part of the paper draws a set of computational implications from the material presented in Part 1. This is done within the general framework of a process model of recognition. Part 2 contains a discussion of the basic types of mechanisms needed for the classes of indirect speech act identified in Part 1. The discussion includes an examination of the dependencies between processes and an initial categorization of the types of knowledge that must be considered in interpreting indirect speech acts.



Toward a Computational Theory of Indirect Speech Acts

Gretchen P. Brown

September 1979

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research under Contract Number N00014-75-C-0661.

Massachusetts Institute of Technology

Laboratory for Computer Science

Cambridge Massachusetts 02139

Abstract

The variety of surface forms that may be used to convey a given speech act pose a major problem in modelling task-oriented (and other) dialogues. Many such forms are so-called indirect speech acts, that is, surface form does not correspond to the (or one) intended speech act. While this topic has received extensive attention from linguists, their concerns have not usually been computationally motivated. In this paper, I present a non-computational analysis of indirect speech act forms with an eye to computational considerations.

The paper is divided into two parts. Part I presents categories and rules for indirect speech acts, justified where possible by traditional linguistic arguments. The second part of the paper draws a set of computational implications from the material presented in Part I. This is done within the general framework of a process model of recognition. Part 2 contains a discussion of the basic types of mechanisms needed for the classes of indirect speech act identified in Part I. The discussion includes an examination of the dependencies between processes and an initial categorization of the types of knowledge that must be considered in interpreting indirect speech acts.

Keywords:

indirect speech acts	speech acts	English dialogue
conversational maxims	semantics	natural language
computational linguistics	pragmatics	language recognition

Toward a Computational Theory of Indirect Speech Acts

Contents

1 Introduction 9

1.1 The Area of Interest 9

1.2 Excluded Topics 13

1.3 The Organization of the Paper 16

PART I: THE TAXONOMY OF INDIRECT SPEECH ACTS PROPOSED 21

2 Defining and Organizing Speech Acts 21

2.1 Some Major Points from Speech Act Theory 21

2.2 Half of an Approach to Taxonomy 24

2.3 The Other Half 27

2.4 The Notion of a Precondition 31

3 The Speech Acts Used in the Examples 35

3.1 Information-Centered Speech Acts 35

3.2 Action-Centered Speech Acts 39

4 Representing Procedures: The OWL-1 Method 49

5 Three Approaches to the Treatment of Indirect Speech Acts 63

5.1 Gordon and Lakoff 63

5.2 Sadock 64

5.3 Searle 66

6 Some General Rules for Indirect Speech Acts 71

6.1 ISA Forms Related to the Maxim of Necessity 72

6.2 ISA Forms Related to the Maxim of Desirability 74

PRECEDING PAGE BLANK-NOT FILMED

6.3 ISA Forms Related to the Maxim of Possibility. 76

6.3.1 Forms Based on Preconditions of Individual Speech Acts. 76

6.3.2 ISA Forms Based on Semantic Cases of Actions. 82

6.4 Tense and Mood Variations. 84

6.5 Examples that Fall Outside the Rules. 87

7 A Taxonomy of Indirect Speech Act Forms. 89

7.1 Overview. 89

7.2 Direct Indirect and Tag Forms. 90

7.3 The Speech Acts Conveyed. 93

7.4 Frozen ISA Forms. 96

7.5 Explicit and Implicit Actions. 100

7.5.1 Must you smoke?. 101

7.5.2 Can you reach the salt?. 102

7.5.3 It's cold in here.. 107

7.5.4 Special Properties of Type 3 Implicit-Action Forms. 110

7.6 Summary: Implications for the General ISA Rules. 112

PART 2: COMPUTATIONAL IMPLICATIONS 117

8 Computational Perspectives on ISAs. 117

8.1 Drawing Computational Implications from the ISA Categorization. 117

8.2 A Look at Related Work in Dialogue Modelling. 119

9 A Framework for Recognition Processing. 123

9.1 A System Configuration for Recognition. 123

9.2 The Event Tree and Recognition Mode. 129

10 General Observations about ISA Recognition. 135

10.1 The Role of the Recognition Framework. 135

10.2 ISAs and the Event Tree. 137

10.3 Some Remarks about ISA Ambiguity. 142

10.4 Dependencies in ISA Disambiguation. 146

11 Outline of a Mechanism for ISA Recognition. 149

11.1 Applying the Rules for Explicit-Action Forms. 149

11.2 Recognizing Frozen ISA Forms. 158

11.3 Recognizing Implicit-Action ISAs. 160

11.3.1 Integrating Implicit Actions into the ISA Mechanism. 161

11.3.2 Recognizing Utterances with Tertiary Forces. 163

12 Knowledge Sources for ISA Recognition. 167

12.1 Domains of Discourse. 168

12.2 Complementary Roles of P1 and P2. 170

12.3 Properties of P1. 172

12.4 Properties of P2. 172

12.5 Exception Conditions. 173

12.6 The Course of the Interaction to Date. 174

13 In Conclusion. 179

References. 183
Appendix A. Core Dialogue Methods. 187
Appendix B. Matching Semantic Representations against Rules. 189
Appendix C. ISA Examples for Rules 6.1 to 6.9. 193
Appendix D: Relating Rule 7.6 to OWL-1 Methods. 201
Appendix E: A Summary of Representation Levels in the General Recognition Model. . 205

1. Introduction

1.1. The Area of Interest

The variety of surface forms that may be used to convey a given speech act pose a major problem in modelling dialogue.[1] Consider, for example, some of the different ways to ask someone else to write a computer program for you:

1.1 I request that you write a program to manipulate blocks for me.
1.2 Write a program to manipulate blocks for me.
1.3 Would/Will you write a program to manipulate blocks for me?
1.4 Could/Can you write a program to manipulate blocks for me?
1.5 I want you to write a program to manipulate blocks for me.
1.6 I want a program to manipulate blocks written.
1.7 I want a program to manipulate blocks.
1.8 I would like you to write a program to manipulate blocks for me.
1.9 I would like a program to manipulate blocks written.
1.10 I would like a program to manipulate blocks.
1.11 Give me a program to manipulate blocks.
1.12 Would/Will you give me a program to manipulate blocks?
1.13 Could/Can you give me a program to manipulate blocks?
1.14 I want you to give me a program to manipulate blocks.
1.15 I would like you to give me a program to manipulate blocks.
1.16 Write me a program, would/will you?
1.17 Write me a program, could/can you?

---

1. I am grateful to William A. Martin and members of the Knowledge Based Systems Group at the M.I.T. Laboratory for Computer Science. Their work provided the foundation on which this paper is based. The OWL-1 knowledge representation formalism used for the process model discussed in Part 2 and the approach to interpreting procedural representations are products of the efforts of this group. I also wish to thank Bill Mark, Candy Sidner, Peter Szolovits, and Jim Weiner for helpful comments on drafts of this paper and Ashok Malhotra, William Mann, and Louis Melli for making dialogue transcripts available to me.

PRECEDING PAGE BLANK-NOT FILMED

Utterances 13 to 117 are commonly called *indirect speech acts*, since the surface form does not correspond directly to the intended speech act. This is not a complete list but is intended only to illustrate the variety of forms that are possible. Similar lists could be made for other speech acts. Requests seem to have the largest number of possible forms, but warnings, suggestions, and questions also appear in a number of variations. In fact, all speech acts with the exception of some ceremonial acts seem to have at least some indirect forms.

For a theory to adequately deal with indirect speech acts, (henceforth, ISAs[2]) it must contain at least the following components:

1. A characterization of the forms that may be used to convey individual speech acts
2. A description of mechanism(s) to relate surface forms to underlying speech acts[3]
3. An account of the conversational implications of choosing a particular surface form

Parts 1 and 2 of this paper relate, respectively, to the first two of these goals. In Part 1 I present a group of distinctions intended as the core of a taxonomy of ISA forms (Goal 1). The categories proposed draw on insights from, but not necessarily the theoretical framework of, previous work in this area (e.g., Sadock, Searle, and Gordon and Lakoff). The categories differ in important ways, however, from existing proposals. Although the

---

2. I hope that the similarity between this abbreviation of *indirect speech act* and IS-A, a name used commonly in the Artificial Intelligence literature for a hierarchic semantic relationship, is not too distracting. The letters of ISA are pronounced individually, not as an acronym.

3. Here and elsewhere in the paper, I use *surface form* to mean simply a written or spoken English utterance.

choice of these categories was strongly motivated by computational considerations, they are supported by independent linguistic evidence as well. This use of non-computational evidence to justify computationally attractive categories is a methodological strength of the presentation.

Part 2 of this paper focusses on the second area cited above. There I concentrate on recognition and describe aspects of an ISA process model. ISA recognition cannot be completely divorced from general issues of recognition, and I have set the process model within a larger framework. ISA processing can, however, be viewed independent of a particular implementation, and that has been done here.

The third area cited above, accounting for the conversational implications of choices, will not be given a systematic treatment here due to the scope of the problem. Some work in this area has already been done, e.g., that by Lakoff [15] and Davison [7].

My account of ISAs will be restricted to task-oriented dialogue, which was defined by Grosz as "one in which two (or more) people communicate for the sole purpose of completing some task" [8].[4] I would extend this definition beyond people to include communicating systems (i.e. computers). A good model of task-oriented dialogue would have obvious practical advantages in the realm of human-machine interaction. Beyond this, task-oriented dialogue has some methodological advantages. First, the task being carried out not only gives structure to the dialogue but also gives valuable independent information to the outsider trying to interpret dialogue transcripts. Information about the

4. Although this definition opens up the possiblity of more than two participants, I will restrict my attention to two-party conversations.

task can be particularly useful in ascribing goals and motivations to dialogue participants. Second, the conversational implications of choosing a particular ISA form (Goal 3, above) are generally less crucial to task-oriented dialogue than, say, to the therapeutic dialogue analyzed by Labov and Fanshel [14]. This does not mean that the implications of choices can be ignored in an ISA model, but it does mean that work on modelling task-oriented dialogue can proceed without waiting for a full systematic treatment of the implications of choices.

In developing this treatment of ISAs, I have concentrated on the areas of similarity between typewritten and spoken utterances, which means that I will have little to say about features unique to speech, e.g., intonation. Certainly more work on the place of special speech mechanisms is warranted. The commonalities between written and oral modes are a good place to start, however, because there is a good deal of lexical and grammatical variation to be accounted for in ISAs.

Two additional remarks: first, examples are drawn from several different sources. A number of examples are drawn from the ISA literature, and some have been taken from transcripts of both typewritten and spoken dialogues. The rest of the examples are manufactured ones, to avoid the necessity of long scene-setting explanations and to more easily highlight the phenomena that are of interest. Although I have been able to find counterparts for some of the manufactured examples in dialogue transcripts, this has not been the case for all of them. In general, I leave the empirical justification of the proposed ISA framework to further research and appeal instead to the linguistic intuitions of the reader.

The second remark relates to representation. A major focus of this paper is on rules and pieces of semantic representation that together, I claim, can be shown to account for a large variety of ISA forms. The rules and semantic representations are presented here in informal English, although they have been translated into a knowledge representation scheme called OWL-1. I feel that semantic representations in general are still in their infancy; we still have only an imperfect idea about what elements should be included in a semantic representation and why. Because of this, little would be gained by the use of the notation here. The loss of precision in using English is hopefully offset by the greater intelligibility of the presentation and the obvious applicability of the ISA rules across representation schemes.

I am, then, considering ISAs in task-oriented dialogue, and I will be working primarily with the commonalities that exist between written and spoken forms. Before a more detailed survey of what the paper will include, I will list some of the areas that *will not* be considered.

1.2. Excluded Topics

There are two sets of topics that will be excluded from consideration. The first relates to phenomena that I do not consider to be ISAs. The second set of topics relates to legitimate ISAs that will be ignored in an effort to limit the scope of the paper.

We start with the non-ISAs. ISAs will be categorized by their intended underlying speech acts, so this excludes speech acts that are not intended by the speaker/sender, P1.[5]

5. To simplify discussion of dialogue exchanges, I will call the speaker/sender of the first utterance considered P1 and the hearer/receiver P2. If subsequent related utterances are

For example, consider the case where P1 intends a simple description of a personal wish but, on hearing P1's utterance, P2 treats it as a request and sets about to make the wish a reality. Whatever the result of P1's utterance, if P1 did not intend the utterance as a request, then I will not consider it one.

Another sort of utterance that will not be considered to be an ISA is one where there is an additional implicit sense arising only from a particular cooccurence of dialogue steps. For example, in a computer console session environment, users sometimes type "Thank you" in a place where other users type "Thank you" followed by "Good-bye". I do not consider "Thank you", when it occurs alone in such a situation, to be an indirect closing. Instead, I consider the closing to be an optional step, which may be omitted in the presence of utterances that uniquely identify the place in the dialogue. A similar optional step would be a statement acknowledgment such as "I understand," "Uh huh" or "O.K." If the next utterance is one that shows that the statement has been understood, an acknowledgment is not strictly necessary. The criterion I am using to distinguish ISAs from utterances preceded or followed by omitted optional steps is that *ISA forms are derivable from conditions associated with the conveyed speech act(s)*. Utterances that "imply" omitted steps do so based on relationships at a level of dialogue structure more aggregate than speech acts.[6] Consequently, just because one utterance U1 implies another, U2, this does not mean that U1

---

discussed, then P1 and P2 continue to refer to the same participants. The only places that I depart from this convention is in reporting on other theories. In these cases, I follow the terminology used by the author of the theory.

6. See Section 4 for a discussion of aspects of this level of dialogue structure. I suspect that application of the framework described there could lead to good progress in characterizing the conditions that permit omitted steps.

should automatically be classed as an ISA with the force conveyed by U2 as an underlying speech act.

This takes care of forms that will not be considered ISAs, and now I briefly list phenomena that, although interesting, are beyond the scope of this paper. The first of the excluded phenomena are cases in which rules or conventions are violated. Examples are deliberate violations such as insincerity and accidental ones such as mistakes. In addition, I will not discuss what I consider to be second order uses of speech acts, such as sarcasm, jokes, or failure to make standard choices (e.g., P1 makes an utterance and has not decided whether it is a question or a request for a nonverbal action).[7] Finally, I have excluded one particular class of ISAs, those identified by Fraser as hedged performatives [9]. This class of ISAs, although extremely interesting, seems to be of relatively minor importance for the type of dialogue I am considering.

A final area that is considered only partially is context. Linguistic and situational context plays a major role in both the generation and the understanding of ISA forms. The role of context clearly cannot be ignored. At the same time, a full account of the nature of context is a massive undertaking, and one that has implications far beyond ISA forms. This paper, then, contains less than the full story on context, although I have given it a relatively large amount of attention. See in particular Section 12.

---

7. One interesting second order use of ISAs is in what can be called *force shift*. In these cases, one speech act form is used to "masquerade" as another. For example, one may use a suggestion form such as "How about picking up the blocks now?" in an environment where authority and role relationships make it clear that the utterance is functioning as a command. In general, force shift seems to be used to give P1 the appearance of greater benevolence or to save face for P2.

1.3. The Organization of the Paper

I will finish this introduction with a brief description of the organization of the paper. Part I begins with a description of the major points of speech act theory that have affected the model presented here. In Section 2 I am particularly concerned with issues of taxonomy, since an account of ISAs depends on clear boundaries between classes of speech acts. Section 3 gives a detailed view of the speech acts used as examples throughout the paper, and Section 4 introduces the OWL-1 method, a structure for representing actions. The preliminaries aside, Sections 5 through 7 focus on ISAs. Section 5 discusses three different theories of ISAs, and the next two sections present a theory that builds on all three of these, while also accounting for additional phenomena. In particular, Section 6 presents a set of general ISA rules. Section 7 shows how these rules must be restricted and augmented to account for the ISA data. This section identifies a set of categories that account for some of the major differences found among ISA forms.

The computational implications of the ISA taxonomy are developed in Part 2. The central thesis of Part 2 is as follows:

> The phenomenon of indirect speech acts is too complex to admit to a single, uniform computational treatment. It is necessary, then, to identify classes of indirect speech acts that share similar computational properties and use different representations and processing strategies for the different classes.

By no coincidence at all the candidates for differing computational treatment are exactly the categories developed in Part 1.

While recognition of ISAs has its own unique aspects, it cannot be viewed entirely

separately from the whole problem of recognizing utterances. For this reason, after some additional introductory remarks in Section 8, Section 9 discusses the points in the recognition process that have the most direct relevance to the processing of ISAs. Section 10 presents general observations on ISA processing in the light of the recognition framework that was described, and Section 11 contains specific proposals for a mechanism. Finally, Section 12 is an initial characterization of the knowledge sources that come into play in framing and interpreting ISAs.

PART I

THE TAXONOMY OF INDIRECT SPEECH ACTS PROPOSED

PRECEDING PAGE BLANK-NOT FILMED

2. Defining and Organizing Speech Acts

2.1. Some Major Points from Speech Act Theory

Any treatment of ISAs necessarily rests on a general theory of speech acts. Fortunately for the theory, but unfortunately for would-be summarizers, the speech act literature is extensive. Since I cannot hope to adequately summarize it here, I will mention only those aspects of the theory that are reflected directly in the account of ISAs to be given. Readers unfamiliar with speech act theory are referred to Austin [2], Searle [29], and, for a wide-ranging bibliography, Verschueren [34]. More specialized references will be cited in the course of this section.

The central thesis of speech act theory is that a strict dichotomy between speech and action is untenable. Before Austin, much of the philosophical interest in language centered around its descriptive uses and the conditions under which statements could be judged right or wrong. While acknowledging this *constative* use of language, Austin identified a class of utterances which he called *performatives*. Standard examples of explicit performative utterances are:

2.1 I name this ship the *Queen Elizabeth*.
2.2 I bet you sixpence it will rain tomorrow.

Once 2.1 is uttered under the proper circumstances, the ship officially has a name, and, once 2.2 is uttered, a bet has been proposed. The important point about these and other performatives is that *saying something* is also *doing something*.

Austin's central insight that language is action was taken up and developed further in

PRECEDING PAGE BLANK-NOT FILMED

the work of Searle. In [29] Searle identifies five types of *constitutive rules* for speech acts. Constitutive rules create or define new forms of behavior, as opposed to regulating or sanctioning behavior. Searle's five types of rules, when taken together, are intended to give necessary and sufficient conditions for performance of a speech act. The five types of rules are:

1. Rules shared by all speech acts are those based on *normal input and output conditions*, which Searle describes as "the range of conditions under which any kind of serious and literal linguistic communication is possible."[8]

2. *Propositional content rules* describe constraints on the propositional content of speech acts. The propositional content of a speech act is, very informally, what the speech act is "about." For example, the propositional content of a request is "a future act A of H." (H is Searle's abbreviation for hearer. S, which appears below, stands for speaker.)

3. *Sincerity rules* are those based on conditions that must occur for the speech act to be performed sincerely. The sincerity rule for request is that the request is uttered only if "S wants H to do A."

4. *Preparatory rules* are those based on conditions that must obtain initially for the speech act to be performed successfully. An example for request is that the request is uttered only if "H is able to do A."

5. *Essential rules* are based on what Searle sees as the essential feature of the speech act, one that distinguishes it from other similar speech acts. For requests, the essential rule is that the request "counts as an attempt to get H to do A."

Throughout the work on speech acts, there is a basic tension between underlying speech acts and the words used to carry out or describe them, commonly called speech act

---

8. [29],p.57.

verbs (henceforth SAVs). Verbs such as *ask*, *request*, *congratulate*, *promise*, *apologize*, etc. can be used to carry out the action conveyed, while other verbs such as *convince*, *persuade*, and *intimidate* are used descriptively.[9] Austin offers a preliminary taxonomy not of speech acts but of SAVs. Searle, who presents a taxonomy of speech acts (see [31]), points out that different SAVs do not necessarily convey different speech acts; some SAVs, for example, differ only in the manner of carrying out the speech act. Searle's example, which is somewhat controversial, is *announce*, since one can announce orders, promises, and reports. Another example might be the use of *remark* and *emphasize* to reflect differences in delivery of, potentially, the same information.

What are the implications of these ideas for a computational theory? First, the view that speech is action means that, as a minimum, speech acts should to be represented in a way that reflects the commonalities with other sorts of actions. At the same time, the representation must be able to reflect the special properties of speech (and written communication) that distinguish speech acts from other sorts of actions. Finally, we want to maintain a distinction between speech acts and SAVs as a source of information about the boundaries of speech acts. All of these themes will play an important role in the treatment of ISA forms presented here.

---

9. This latter group of verbs focusses not on the speech act (the *illocutionary act*), but on the effects of the speech act on the hearer. Austin calls the acts conveyed by such verbs *perlocutionary acts*. It seems to me to be desirable, however, to emphasize the commonality of the speech act between, say *argue* and *convince*. These refer to the same type of act, but for the latter, the effect is described as well. See [6] for a similar treatment of perlocutionary acts within a computational framework.

2.2. Half of an Approach to Taxonomy

Austin estimated the number of SAVs to be between one and ten thousand. Assuming that the number of speech acts is also of this order, then clearly a taxonomy is essential to a general theory of speech acts. It turns out that a taxonomy, or at least a principled approach to one, is also crucial to a treatment of ISA forms. If one wishes to claim that forms associated with one speech act are used to convey another, then one must be quite sure that the two speech acts are really distinct. Otherwise, the "indirection" is no indirection at all.

In this subsection I present an approach to taxonomy that combines aspects of earlier work with some new additions. Before plunging in, I wish to emphasize that this is not seen as the only speech act organization possible, nor is it seen as the only computationally useful one. It is, however, one that plays a significant role in ISA organization. Section 3 presents additional speech act categories that will be used in the ISA analysis.

For the basic approach to taxonomy, I take a high level set of categories from Searle but use a defining criterion derived from work by Verschueren. At the most general level of speech act categories, I start with the five classes identified by Searle [31]: representatives, directives, commissives, expressives, and declarations. Searle bases these five classes on the "illocutionary point" of the act, that is, on "differences in the point (or purpose) of the (type of) act." Because of the importance of these speech act classes, I will give Searle's definition of each, excerpted from [31].

> Representatives (e.g., asserting): The point or purpose of the members of the representative class is to commit the speaker (in varying degrees) to something's being the case, to the truth of the expressed proposition.
>
> Directives (e.g., requesting): The illocutionary point of these consists in that fact that they are attempts (of varying degrees and hence more precisely, they are determinates of the determinable which includes attempting) by the speaker to get the hearer to do something.
>
> Commissives (e.g., promising): Commissives...are those illocutionary acts whose point is to commit the speaker (again in varying degrees) to some future course of action.
>
> Expressives (e.g., thanking): The illocutionary point of this class is to express the psychological state specified in the sincerity condition about a state of affairs specified in the propositional content.
>
> Declarations (e.g., nominating a candidate) It is the defining characteristic of this class that the successful performance of one of its members brings about the correspondence between the propositional content and reality...

These five classes give an excellent start toward a computationally useful taxonomy of ISAs. The notion of illocutionary point, however, has been criticized by Verschueren [35] for its combination of illocutionary (i.e., act-related) and perlocutionary(i.e., effect-related) properties. (The definition of directives mentions the hearer explicitly, but other categories do not.) Verschueren views SAVs as reflecting the speaker's intention to bring about some effect in the hearer by pronouncing an utterance.

Verschueren's focus on the hearer can be profitably applied to speech acts as well as SAVs; the one addition that I would make is that some speech acts can have more than one intended effect on the hearer, so that the classification criterion must be according to the

principal effect that the speaker (P1) intends to have on the hearer (P2). Using this approach, Searle's five categories can be redefined as follows:[10]

> Representatives have as their principal intended effect that P2 come to know that P1 believes something. [11]
>
> Directives have as their principal intended effect that P2 take responsibility for carrying out some action.[12]
>
> Commissives have as their principal intended effect that P2 accept P1's commitment to take responsibility for carrying out some action.
>
> Expressives have as their principal intended effect that P2 come to know that P1 feels something.
>
> Declarations have as their principal intended effect that P2, and possibly society as well, come to perceive and accept a change in reality.

We have, then, Searle's speech act categories and a defining criterion based on a suggestion by Verschueren. The criterion is suited to a computational approach because it points in the direction of larger patterns of action. It takes into account both participants in a dialogue and foreshadows responses to the speech act (i.e. whether the speaker's intention

---

10. The first three criteria are closely patterned after similar ones given by Verschueren in [35].

11. I use "come to know" rather than "know" here to indicate that I am defining the principal intended effect as an action rather than simply as a state. Expressives, commissives, and declarations are given the same treatment.

12. I use *take responsibility* here to indicate that P2 may either carry out the action or subcontract some or all of it.

is satisfied, is not satisfied, or whether the response falls somewhere in between). I will come back to these issues in Section 4.

This takes care of the top of the speech act taxonomy. Besides its use in forming the most general categories, the "principal intended effect" criterion can also be used to form intermediate levels of speech act classes. Although I have found such categories useful for other purposes, they are not necessary for the ISA analysis. I merely mention their possibility and go on in the next subsection to the other half of the taxonomy, the question of how to distinguish speech acts from each other.

2.3. The Other Half

This brings us to the level of speech acts.[13] Despite all the work that has been done in speech act theory, I believe that we still lack a totally satisfying answer to the question of how one speech act is to be distinguished from another. Searle's constitutive rules give a way to distinguish speech acts from other sorts of actions and give a framework for *expressing* distinctions between speech acts, but they do not specify exactly where lines are to be drawn between different speech acts. The "principal intended effect" criterion gives a necessary condition for differentiating between speech acts, but it does not seem to be sufficient for the purposes of studying ISAs. For example, warn and command are generally considered to be different speech acts, but both can be viewed as attempts by P1 to get P2 to take responsibility for carrying out some action. Granted, warnings as a class have the

---

13. I will refer to generic speech acts as, simply, speech acts. Realizations of a generic speech act will be referred to as particular speech acts.

extra restriction that the action is to avoid some condition that is perceived by P1 to be against P2's best interests. There is nothing to stop one, however, from commanding the same sort of action. Thus, an utterance such as 2.3, might be construed as a warning or a command, depending on whether authority or "good sense" is being appealed to.

2.3 Keep off the ice.

At such borderline cases, the difference seems to rest on the nature of P2's *obligation* to take responsibility for the action. Warnings seem to be empowered by an appeal to P2 to act reasonably, i.e. in P2's own interest. Commands are empowered by the authority relationship that exists between P1 and P2. If command and warn must be distinguished based on the nature of the relationship appealed to, then the "principal intended effect" criterion is clearly not sufficient. I need, then, a principled approach for the speech act level of the taxonomy.

I start the search for a criterion with the conviction that the resulting set of speech acts should look familiar. They should not be radically different from the categories found in the literature and, more basically, those reflected in English and other languages. The resulting system should allow the use of familiar terms such as *warn, command, request, ask, tell*, etc.

After classifying a number of speech acts, I have found no one simple criterion that yields the familiar set of speech acts. There are, however, three guidelines for what constitutes a speech act that I have found useful:

1. SAVs suggest a minimal group of speech acts, if one omits:

a. general SAVs such as *say*, which can refer to more than one speech act (where "speech act" is delineated by application of some other guideline).

b. context-marked SAVs such as *reply*, *interject*, and *answer*, which mark the place of the utterance in the discourse.[14]

c. aggregate SAVs such as *describe* which refer to more than one speech act acting as a unit.

d. SAVs that differ only according to PI's view of the importance of the speech act, e.g., *mention* used for stating.

2. Sets of linguistic realizations that are substitutable for each other without seriously disturbing the flow of a dialogue should have a separate speech act with which they can be associated.

3. Since particular speech acts occur as part of larger linguistic and non-linguistic *activities*, speech acts should be chosen so that they are easily integrated into these larger patterns of action.

No matter how refined these guidelines become, they probably cannot be made to function independently. That is, no one of them can be turned into a criterion that is sufficient to clearly discriminate between all the speech acts that one would expect. With respect to SAVs, at least in English there are candidates for the status of speech act that have no corresponding SAV. For example, there is a separate set of surface forms for giving instructions, and so, according to guideline 2, one would want to identify a separate speech act. There seems to be no English SAV that expresses the notion of giving one-

14. In barring the referents of context-marked SAVs from independent speech act status I differ with Searle [31]. The difference is not particularly serious and means only that the class of independent speech acts that I am admitting is slightly smaller than his.

utterance-worth of instructions. (*Instruct* is either a request or an aggregate.) SAVs, then, provide only a minimal guideline.

The difficulty with Guideline 2 comes in narrowing down the notion of "seriously disturbing the flow of dialogue," since probably no utterance forms are ever completely synonymous. The choice of one form over another seems to always involve some implications, e.g., politeness, formality, combativeness, etc. There seems to be a grey area between characteristics of utterances that should be accounted for by "implications of choice" and those that should be accounted for by assuming a difference in conveyed speech act.

With respect to the third guideline, "ease of integration" is almost certainly not a simple property. The process of representing larger patterns of action is open to influences beyond those local to the speech act. Frequently, altogether different sets of speech acts can be supported if different choices are made at other places in the representation. (For further discussion of larger patterns of action, see Section 4.)

Even though these guidelines cannot function independently and in some cases they will still leave some grey areas, they can still have important implications for a computational theory. A set of speech acts that follows the guidelines will have the following properties:

1. The speech acts will be easily related to English SAVs.

2. The speech act representations will provide a useful place to associate information about groups of linguistic realizations.

3. The speech acts will fit into larger patterns of action, allowing us to model not only isolated speech acts but also speech acts as they occur in dialogue.

The speech acts used in rules or examples in this paper are introduced in Section 3. They represent an application the three guidelines given here. After working with them for some time, I believe that they display the three advantages listed above. These advantages begin to be apparent in the discussion of the computational model in Part 2.

2.4. The Notion of a Precondition

In defining speech acts with ISAs in mind, it is useful to shift the emphasis slightly in the treatment of constitutive rules. I will therefore introduce the notion of a *precondition* of a speech act. Preconditions are a combination of Searle's sincerity conditions and preparatory conditions viewed *from the point of view of PI*[15] (where PI is the agent of the speech act). Because PI's point of view may not correspond to reality, particular speech acts may meet all preconditions and still be defective in the sense used in Searle [29]. The preconditions that I am interested in are those that differ from one speech act to another. For this reason, I will ignore general requirements such as the requirement that a given speech act must be intended by PI.

Although the actual representation that I am assuming for the preconditions is OWL-I, for the purposes of this paper I present the preconditions in informal English. Despite the informality, some words are used in a restricted sense and several simple conventions have been followed. I will discuss these conventions and some of the restrictions here, and the rest of the restrictions will be explained as examples are presented.

---

15. Searle's essential conditions are not directly reflected in preconditions, although they seem to be derivable from them. I see essential conditions as an amalgam of several different types of information, some of which would be explicit in preconditions and the rest of which would probably be implicit.

The first convention is that P1 in preconditions refers to the speaker/writer of the speech act and P2 refers to the hearer/reader. Speech act conditions centered on intentions of P1 (in the philosophical sense of *intention*) will have preconditions that correspond directly, i.e.,

P1 <intention> <object of intention>

e.g., P1 wants some action done.

All other speech act conditions are recast with respect to P1's model of the world, i.e.,

P1 believes <speech act condition>

e.g., P1 believes that P2 can take responsibility for a (particular) action.

This takes care of the conventions; two restrictions are also of interest, *want* and *believe*. For both types of precondition, I have represented goals of participants by the verb *want*, i.e., "P1 wants <state or action>" or "P1 believes that P2 wants <state or action>". *Believe* is more complicated. As used in preconditions it is broader than *know*; that is, the direct object of *believe* may be either a fact or an opinion. By "P1 believes X" I mean that P1 has some world model that contains X and P1 believes that X corresponds to an aspect of reality. By "correspond to an aspect of reality" I mean that actions based on X reach the expected results or, if they do not, the deviation from the expectation is not caused by X. As the philosophical literature attests, much more remains to be said about facts, opinions, and reality. To avoid infinite -- or at least lengthy -- regress, I will cut off the discussion here and appeal to the reader's intuitions.

The notion of a precondition is operationally useful in a computational theory. The

preconditions of a speech act are groups of conditions that are expected to obtain before the speech act is initiated. As such, they should be intuitively obvious to speakers of the language. One pair of linguistic tests for "preconditionhood" relate to the way that P2 may claim that a failure has occurred in P1's speech act. For speech acts centered on intentions of P1, P2 may reply with form 2.4. For others, form 2.5 is possible.

2.4 You don't really <intention>.
e.g., You don't really want me to do that.

2.5 What makes you think <direct object of initial "P1 believes">?
e.g., What makes you think I can tell you?

Note that these tests are necessary but not sufficient to distinguish the preconditions particular to a speech act. They also apply to conditions associated with speech acts in general. (*See Section 6.3 for some of these general conditions.*)

Several examples of the sets of preconditions associated with different speech acts are found in the next section, and the usefulness of the concept of a precondition for ISA analysis will become apparent in Section 6.

3. The Speech Acts Used in the Examples

In this section we take a closer look at four commonly occurring speech acts. These will be used as examples throughout the paper. The examples are drawn from two of Searle's categories that are particularly relevant for task-oriented dialogue: directives and commissives. (Representatives, also very important in task-oriented dialogue, are treated as a class for reasons given below.)

The four speech act examples are presented here in terms of two additional categories, which will be useful later on in the presentation of the ISA categories. (Recall the warning that Searle's five categories were seen as one of *several* useful divisions of speech acts.) The split that I am making is on types of propositional content. The two speech act classes will be called *information-centered* and *action-centered speech acts*; they are defined in the two subsections below. Note that the two categories do not cover all the speech acts in Searle's five categories. While it would be possible to extend the alternative categorization scheme, this is not necessary for the purposes of this paper.

3.1. Information-Centered Speech Acts

Information-centered speech acts are those whose intended effect is the transfer of information between P1 and P2 with no additional participants involved. Information-centered speech acts include Searle's representatives and other types of speech acts such as the directive ASK.[16] For the purposes of this paper, I will need only ASK. While

16. Here and elsewhere, I will capitalize the names of speech acts that are used to refer to the speech act definitions given in this section. Small letters will be used for references to other definitions.

PRECEDING PAGE BLANK-NOT FILMED

representatives will play an important role in many aspects of the ISA analysis, I will not be focussing on ISAs of individual representatives. The process of drawing distinctions between individual representatives raises many interesting and important philosophical issues, which are, unfortunately, beyond the scope of this paper.

The propositional content and the preconditions of ASK are stated informally in Figure 3.1 We have said that ASK is a directive, i.e., an attempt by P1 to get P2 to take responsibility for carrying out some action. In this case, the action desired is that P2 give P1 some information, the answer to a question.[17] As I have defined ASK, it requires that P1 want the answer (precondition 1). To be consistent, P1 must also believe that he or she does not know the answer; this is part of the definition of *want*. The first precondition, then, means that certain types of questions are not classed under ASK. Among these are test questions and ceremonial questions where P1 already knows the answer.

The preconditions in Figure 3.1 should be self-explanatory, with the possible exception of (V). The notion of obligation used here is a more specific version of the generalized obligation that Labov and Fanshel use for requests in [14]. (I am using the same set of obligations for REQUEST as for ASK; further discussion of the relationship between these two speech acts is given below.) An example of an obligation arising from complementary roles would be the obligation that applies in an interviewer/interviewee relationship. Authority obligations are slightly more difficult to identify, since, especially in contemporary American society, most authority arises from roles. Still, it is possible to identify authority

17. I appeal to intuitive notions of *question* and *answer* here, although much more can -- and eventually should -- be said. It appears that a specification of the semantic structure of questions and a full treatment of the semantics of answers will be sizable tasks.

Figure 3.1. The preconditions for ASK

propositional content: some question

I. **PI wants to know an answer to the question.**

II. **PI believes that P2 can tell an answer to the question.**

III. **PI believes that P2 is willing to tell an answer to the question.**

IV. **PI wants P2 to tell PI an answer to the question.**

V. **PI believes that P2 has some obligation (a role obligation, authority obligation, or general obligation to be helpful) to PI to tell PI an answer to the question.**

Comments

(I) *Want* **here and in (IV) implies that PI does not already know the answer. The case where PI does know and merely wants to know if P2 knows (and where P2 knows that PI knows) is classed as a different speech act.**

(I) *Know* **is considered to be a restricted form of** *believe*; **while anything can be believed, only facts can be known. (For ASK, the "fact" is that some proposition is the answer to the question.)** *Fact*, **of course, needs its own definition which I do not have space to go into here; instead, I appeal to intuition and observe that facts are most strongly contrasted with opinions.**

(II)-(V) *Tell* **is used here to mean "utter a representative."**

(V) More than one of these different types of obligations may apply at once.

obligations that, while they may be carry-overs from roles adopted by the participants, do not arise from current complementary role relationships. For example, if a teacher encounters a student in the halls of the school and the teacher asks a question, the student is obliged to answer by an authority obligation. Certainly, the teacher's authority can be seen as arising from the role, but, in this context, the student need not necessarily be a student of *this* teacher for the obligation to apply. Thus, the authority obligation stems from each participant playing some role, but the role relationship is not complementary in the same way that it is for teachers and students in a classroom situation.

Winding up the discussion of obligations, we come to the "general obligation to be helpful." Being helpful here means that one plays a part in furthering the other individual's goals. Examples of appeals to this obligation would be an ASK to a stranger for the time or an ASK to a friend for the football scores. This type of obligation seems to arise simply from the inequality of knowledge between the participants; if P2 knows something that P1 wants to know, then P2 is obligated to "be helpful" and share the knowledge. This obligation is not absolute (nor are role or authority obligations), since it may be overridden by other obligations.

To finish up the discussion of ASK, note that ASK shows strong similarities to REQUEST, which is discussed in the next subsection. REQUEST is used primarily for nonverbal actions. The distinction between ASK and REQUEST is based on the second and third guidelines from Section 2.3.[18] With respect to Guideline 2, ASK has the

---

18. The first guideline seems fairly inconclusive in this case. True, the SAV *ask* can be used in the following two patterns: "asking someone something" and "asking someone to do something." These two do not, however, correspond completely to a distinction between describing an ASK and describing a REQUEST. For example, "I asked Harry if he would

interrogative as one of its direct forms (what I am calling its *simple form*; see Section 7.2). REQUEST has only the imperative as a comparable form. With respect to Guideline 3, we see differences in the patterns of action in which ASK and REQUEST take part. One example is the following pair of responses:

3.1 I can't.
3.2 I can't tell you.

Here, 3.1 can be a response to a direct (i.e. imperative) REQUEST, but it cannot be used synonymously with 3.2 as a response to a direct (i.e. interrogative) ASK.

All this is not meant to deny that ASK and REQUEST are closely related. Comparison of Figure 3.4, below, with Figure 3.1 shows strong similarities in the preconditions of the two speech acts. These strong similarities do not, however, necessarily justify viewing ASK as a kind of REQUEST. Instead, the similarities can be seen as a result of the common membership of ASK and REQUEST in the directive class.

3.2. Action-Centered Speech Acts

An *action-centered speech act* is defined as a speech act that has as its propositional content an action A1 of a type to be specified where P1 wants A1 to occur. Here, A1 is either non-verbal or a verbal action directed at someone other than P1.[19]

---

lend me a pencil" may be used to report a REQUEST, and "I asked Harry to tell me the answer" may be used to report an ASK.

19. Besides the normal "active" actions I include as actions the maintenance of a state, e.g., being noisy, and the avoidance of an action, e.g., being quiet. The definition of action-centered speech acts is meant to exclude verbal actions such as P1 asking P2 for information but to include P1 asking P2 to give information to a third party, P3. The first case is already covered by ASK, as defined above. I consider the second case to be action-centered

Action-centered speech acts, then, are acts that attempt to *generate* an action, not merely refer to one (as in, for example, thanking someone for some action). Note that the definition above does not specify that the intended effect be the principal one. The implication of this is that action-centered speech acts are not restricted to the directives. The action-centered speech act class also includes members of the commissive and declaration classes where one of the intended effects is to cause the type of action specified above. The action-centered speech act class, then, is formed using "intended effect" as a criterion and cuts across Searle's five major speech act classes.

The action-centered speech acts used as examples in this paper are REQUEST, SUGGEST, and OFFER. I will summarize the major differences between them first and then give the preconditions of each. The first difference is the nature of the action expected of P2. For OFFER, the expected action is one of accepting *another* action. For REQUEST, P2 is expected to take responsibility for carrying out an action desired by P1, and, finally, for SUGGEST P2 is expected to consider a plan of action. A second difference between the three speech acts lies in the obligations invoked. REQUEST depends on appeals to one or more of the three types of obligation discussed for ASK, while OFFER and SUGGEST depend on a combination of self-interest obligations and obligations to "be helpful." The choice of obligation for OFFER and SUGGEST is discussed further below.

---

rather than information-centered because it is much closer to a nonverbal action: P2 must not only evaluate his or her obligation to P3 vis a vis the information (P3's right to know) but P2 must also evaluate his or her obligation to P1 to be the conveyer of the information (P1's right to request). This level of complexity is closer to action-centered than to information-centered speech acts. Note that my argument here is based on Guideline 3. I am appealing to larger patterns of action to draw speech act boundaries, which then percolate up into class boundaries. A case can also be made based on Guideline 2: the simple interrogative form may be used for directives involving verbal actions with P1 as the destination, but it may not be used for verbal actions with a third party as destination.

Taking each of the three example speech acts in turn, I start with OFFER. Although promising is the favorite commissive in the speech act literature, OFFER will be considered here because it has a rich ISA structure. It is also a speech act that is, in a sense, doubly action-centered. An OFFER conveys, first, PI's desire to perform a service for P2 or to give, lend, etc. something to P2. Second, OFFER conveys PI's desire that P2 accept the service, gift, loan, or whatever. The preconditions given for OFFER in Figure 3.2 clearly reflect this split.

The appearance of the "he helpful" obligation in OFFER (VII) needs explaining. The concept that I am trying to express is PI's obligation to accept P2's OFFER. At first glance, it seems enough to say that P2 has a general obligation to act in his or her own self interest (precondition VIII). This obligation, coupled with the fact that the propositional content of an OFFER is an action expected to benefit P2, gives rise to an obligation for P2 to accept the OFFER. There seems to be more than simple pragmatism, however, involved in accepting OFFERs. The obligation to accept is not merely P2's obligation to further his or her own goals, but also P2's obligation to PI to further *PI's* goals (i.e., P2's obligation to be helpful, precondition VII). In accepting an OFFER, P2 is enhancing PI's image as a benevolent person, PI's satisfaction in giving, etc. By accepting, then, P2 is furthering PI's goals and being "helpful."

The next action-centered speech act that we come to is SUGGEST (Figure 3.3). Recall the distinction above that SUGGEST is a directive that P2 *consider* an action, not necessarily that P2 carry it out. Note that this treatment of SUGGEST is supported by the common use of a response such as 3.3.

Figure 3.2. The preconditions for OFFER

propositional content: some action that P1 believes will be of benefit to P2

I. P1 wants to take responsibility for the action.

II. P1 believes that P1 can take responsibility for P1's part of the action.

III. P1 is willing to take responsibility for P1's part of the action.

IV. P1 wants P2 to perform some action that complements P1's part of the action.

V. P1 believes P2 can perform some action that complements P1's part of the action.

VI. P1 believes that P2 would be willing to perform some action that complements P1's part of the action.

VII. P1 believes that P2 has an obligation (to "be helpful") to P1 perform some action that complements P1's part of the action.

VIII. P1 believes that P2 has an obligation to P2 (by virtue of P2's own self-interest) to perform some action that complements P1's part of the action.

Comments

(I)-(III) "Take responsibility" is used here and elsewhere to permit subcontracting. Whether P1 actually performs the action of not, he or she still remains responsible for the results.

(IV)-(VIII) Examples of complementary actions would be (physically) taking food OFFERed by a hostess or getting into a car and sitting in response to an OFFER of a ride from a friend. A general way to refer to P2's performance of a complementary action in response to an OFFERed action is to say that P2 *accepted*, e.g., "Jane thanked Paula and accepted the gift."

Figure 3.3. The preconditions for SUGGEST

propositional content: an action

except for: actions in which P1 and P2 share joint agency

I. P1 wants P2 to consider taking responsibility for the action.

II. P1 believes that P2 can consider taking responsibility for the action.

III. P1 believes that P2 is willing to consider taking responsibility for the action.

IV. P1 believes that P2 has an obligation (to "be helpful") to P1 to consider the action.

V. P1 believes that P2 can take responsibility for the action.

VI. P1 believes that P2 is willing to take responsibility for the action.

VII. P1 believes that there are some reasons why the action is desirable.

VIII. P1 believes that P2 has an obligation to P2 (by virtue of P2's own self-interest) to consider taking responsibility for the action.

Comments

(propositional content) In excluding actions where P1 and P2 share common agency, I am merely arguing for a separate speech act, e.g., SUGGEST-COMMON-ACTION, to cover such cases. I justify this distinction by an appeal to Guideline 2: this second speech act has some forms not shared by simple SUGGEST, e.g., "Let's <action>."

3.3 That's a good idea.

Note that the action to be considered by P2 may be a verbal action directed at P1. For example, to help P2 solve a problem, P1 might SUGGEST 3.4.

3.4 Why don't you tell me what you've done so far.

The unrestricted nature of the action suggested does not, however, threaten the status of SUGGEST as an action-centered speech act. This is because the action in the propositional content is that P2 *consider* the second action, and considering something is a nonverbal action

The obligation preconditions for SUGGEST (IV and VIII) are similar to those for OFFER. Once P2 has acknowledged that an action is desirable (precondition VII), there is a default obligation that P2 consider taking responsibility for carrying out the action (VIII). SUGGEST also appeals to P2's obligation to P1 to help further P1's goals (e.g., enhanced self-esteem and recognition for P1). The obligation arises from the fact that a goal of P1 is involved in a SUGGEST (precondition I) This obligation is generally not as prominent for SUGGEST acts as for OFFERs, but it still plays a role. Consider, for example, the politeness constraints involved in responding to a SUGGEST. A response such as 3.5 violates the obligation in (IV).

3.5 That's a terrible idea.

The third action-centered speech act that will be considered here is REQUEST, the directive by which P1 attempts to get P2 to take responsibility for carrying out some non-

verbal action or verbal action not directed at P1. One other class of actions that I am currently excluding from REQUESTs is remembering previous REQUESTs; that is to say, REQUEST as I define it does not include reminders. This can be justified by an appeal to Guideline 1 due to the difference in SAVs *request* and *remind* as well as by an appeal to Guideline 3. (Reminders are often followed by apologies or thanks by P2, where simple REQUESTs are not.) Figure 3.4 lists the preconditions for REQUEST.

In defining REQUEST, I follow Labov and Fanshel in citing an obligation. Just as for ASK, I further specialize the notion of obligation into authority, complementary role, and "be helpful" obligations.[20] Another point that is worth mentioning is that my treatment of obligation subsumes the notion of P1's right to invoke the obligation. (See [14], p.78) The obligation stated for REQUEST and ASK is a three-place relationship between P1, P2, and the thing that P2 is obliged to do. P1 may be part of a larger set, so that the particular obligation may be to both P1 and the society in general, e.g., the obligation to drive carefully. Given this formulation, P1 has the right to invoke the obligation because P1 is one of the parties to the obligation.

***

This takes care of the speech act examples that will be used in this paper. There is one more piece of foundation that must go into place before the taxonomy of ISAs can be

---

20. For REQUEST and, to a lesser degree for ASK, I have been tempted to define three different speech acts, one for each type of obligation. While there is some Guideline 2 (difference in surface form) evidence for this, the bulk of the argument would have to rest on Guideline 3 (action pattern) evidence. The modelling of dialogue patterns is at too early a stage for Guideline 3 evidence taken alone to be conclusive, so I stay with the more standard treatment of REQUEST.

presented. Accordingly, the next section presents a general representation scheme for actions.

Figure 3.4. The preconditions for REQUEST

propositional content: a non-verbal action or a verbal action whose destination is not P1,

except for: the action of remembering a previous REQUEST

I. P1 wants P2 to take responsibility for carrying out the action.

II. P1 believes that P2 can take responsibility for carrying out the action.

III. P1 believes that P2 is willing to take responsibility for carrying out the action.

IV. P1 believes that P2 is obligated to P1 (and possibly to others) to take responsibility for carrying out the action. (This obligation may be a role obligation, an authority obligation, or a general obligation to be helpful.)

Comments

(I) "Want" implies that P1 believes that the action has not already been carried out.

(III) At least for the purposes of reasoning, "willing" here includes P2 not being grossly unwilling.

(IV) More than one of these different obligations may apply at once.

4. Representing Procedures: The OWL-I Method

In this section I discuss a general representation scheme for actions and patterns of actions. The discussion will add new perspectives to the treatment of taxonomy developed in the last two sections. More important for my purposes here, this section will also form the foundation of the analysis of ISAs that is presented in Sections 6 and 7. Note that the scheme for representing actions described here is not the only one that has been proposed. I defer a discussion of some of the alternatives until Section 8.2, where a comparison of static representations can be combined with a look at the way that the representations are used in language processing.

I start the discussion of patterns of action with OWL-I, which is a formalism for representing knowledge. The basic unit of OWL-I is called a *concept*, and concepts are related to each other by a network of indices. The important point about concepts for this paper is that they can be grouped into larger structures. The structure of interest here is called a *method*.[21] Methods provide a high level, hence relatively declarative, representation for procedural knowledge. The high level nature of the representation means that methods can be used as the basis of an explanation facility (see Swartout [32]). Moreover, this property in combination with the highly interrelated nature of OWL-I concepts means that method structures can be a very important aid in searches of a knowledge base. My discussion of methods will avoid issues of notation. English translations are used for

21. OWL methods were initially defined by William A. Martin. OWL-I is the version of OWL that was used in the dialogue project; OWL has continued to develop since that time, and OWL-II is the current version. For an introduction to OWL-I, see Szolovits et al.[33]

PRECEDING PAGE BLANK-NOT FILMED

examples, and the reader interested in OWL-1 notation is referred to Hawkinson[13] and Brown[3].

With respect to content, methods are used to represent both linguistic and non-linguistic actions. For example, they can be used to represent an action such as writing a computer program for someone else, which normally contains both linguistic and non-linguistic steps. Methods can also be used to model totally non-linguistic activities such as baking a cake.[22] In the examples I have considered and in the rest of this paper, however, I will focus on what will be called *core dialogue methods*. This is a group of semantic domain independent methods built around speech acts. (Core methods may, however, contain some non-linguistic steps, e.g., to represent mental processes.) An example of a core method is request-and-respond, which appears in an English translation in Figure 4.1. (The details of this example will be explained below.) English translations of some other core methods can be found in Appendix A.

Implicit in the preceding discussion is the point that methods may have one participant or several. The core dialogue methods model the parts played by both participants in the dialogue, a property I call *speaker independence*. Methods can therefore be used either from the point of view of a non-participant to model both parts heard (or read) or from the point of view of a participant, in which some utterances are generated and some understood.

Looking at the structure of methods in more detail, we start with the three main parts: a header, argument specifications, and a procedural body. The header is the method's

22. See Long [17] for the use of OWL-1 methods to model programming knowledge.

REQUEST-AND-RESPOND

OBJECT: an action that can be the object of REQUEST
exceptions: 1. a helping action 2. a repetitive action
AGENT: a person or computer system
CO-AGENT: a person or computer system

method:

1. The AGENT REQUESTs the action of the CO-AGENT.
2. The CO-AGENT now knows the action desired.
3. The CO-AGENT checks to see whether beliefs from the preconditions of the REQUEST are justified.
4. The CO-AGENT acknowledges the REQUEST.
5. The CO-AGENT, at the appropriate time, takes responsibility for carrying out the action.
6. PRINCIPAL-RESULT: Any results of the action desired obtain.

recovery path 1: if the CO-AGENT can't do what the AGENT says --
If the REQUEST is based on a role obligation and the action REQUESTed is not in one of the CO-AGENT's roles with respect to the AGENT --
R1.1 The CO-AGENT says the action is not in his or her role.
R1.2 If the CO-AGENT wants to be helpful
and
if the CO-AGENT knows a likely participant to take responsibility for the action desired
then
the CO-AGENT refers the AGENT to this likely participant.
recovery path 2: if the REQUEST was framed in a general way and the CO-AGENT can only do a more specific version --
R2.1 The CO-AGENT describes the specific version that he or she can do.
R2.2 If the more specific description matches the AGENT's goal --
then
the AGENT says that is what he or she wanted
and
REQUEST-AND-RESPOND continues using the new description
otherwise
dialogue failure: the CO-AGENT can't do what the AGENT wants.

Figure 4.1. An English representation of a method for REQUESTing an action and getting a response.

unique name. Argument specifications, organized by semantic cases (see below), are used for type checking of inputs to the method (*input cases*) or to specify the form of results (*output cases*). The procedural body is divided into two parts: (optional) prerequisites and procedure steps. The prerequisites of interest here are those that are states. A stative prerequisite of an action is a condition that must obtain before that action is carried out. If the condition does not hold, then one must bring it about before carrying out the action.

The other part of the procedural body is the set of procedure steps, which come in two basic varieties, *standard path* and *recovery path*. This distinction will be discussed further below, but, basically, standard paths represent the ways that an exchange can "go right," while recovery paths give some of the possible measures to be taken when an exchange gets "off the track." A recovery path may be initiated when an expectation set up by an ongoing method is violated.

Since REQUEST is the speech act from which I will draw most of the examples, I will use the request-and-respond pattern to illustrate the discussion of methods. In Figure 4.1, *request-and-respond* corresponds to the header of the actual OWL-1 method. Request-and-respond takes three input cases -- OBJECT, AGENT, and CO-AGENT -- and one output case, PRINCIPAL-RESULT.[23] AGENT and CO-AGENT are the two participants in the method; by convention, the AGENT of an entire dialogue method is the AGENT of the first step, so that we can identify the AGENT of the method with the participant who starts it off. The OBJECT of request-and-respond is the action desired by the AGENT.

---

23. Semantic case names will be capitalized to avoid confusion with the normal use of these words. There should be no confusion with speech act names, which I am also capitalizing.

[24] Input case specifications typically contain variables, which are like other concepts except that their composition follows special rules. The fact that other concepts may be bound to variables permits representations to be evaluated with respect to some environment. (This is important when methods are used in recognition; see Section 9.) One property of input case specifications omitted from Figure 4.1 is their use as a place to associate further entry conditions, i.e., conditions that must be met before the method is applied. This point is discussed further below.

Note that input cases are associated with methods, not surface English verbs. A list of the major OWL-1 input cases, with informal explanations was written by William A. Martin and is reproduced with minor changes in Figure 4.2. I refer to these as the *major* cases because I do not see Figure 4.2 as an exhaustive list. There will be a significant number of inputs to actions that this list does not cover. Many of these can be fitted into the basic framework as more specialized versions of the major semantic cases. It appears, however, that some actions have totally idiosyncratic inputs which will have to be treated as special cases (in both senses of *case*).

Besides input case specifications, I said that OWL-1 methods may have associated output case specifications, i.e. results. One important notion here is that of ***principal result***, which is the main result of the method and, typically, the reason that the method is

---

24. The OBJECT of request-and-respond excludes helping actions, i.e., those in which the AGENT and CO-AGENT divide responsibility for the action. These are represented in another core method, request-and-help (see Appendix A.) Also excluded are repetitive REQUESTs, e.g., a standing order for publications. More effort must be put into bookkeeping so that the repetitive REQUEST can be responded to repeatedly. The steps differ enough to warrant the use of a separate method.

Figure 4.2. The major OWL-1 input cases

OBJECT

This is the input to the method that is most important to the identity of the process. For example, the process of baking bread is similar no matter who does it, but baking a souffle is a very different process from baking bread.

ex. Make *a statue* from clay.

Empty *the boat* of water.

AGENT

The AGENT is the participant responsible for the action.

ex. *I* hit the ball.

CO-AGENT

This is someone who has a responsibility level equal to that of the AGENT for some part of the action.

ex. *Harry* helped me move the couch.

BENEFICIARY

This is an individual who receives the benefit of the action without major active participation in it.

ex. Answer the phone for *John*.

SOURCE

The SOURCE describes a former position or state of some other case-filler in the action (usually the OBJECT).

ex. Take a block from *the box*.

DESTINATION

The DESTINATION is to the future as the source is to the past. It is a point, object, etc. toward which some other case-filler in the action (usually the OBJECT) is tending.

ex. Tell it to *the judge*.

TRAJECTORY

This is the path taken by some other case-filler in the action.

ex. Run *across the street* to the store.

DIRECTION

This is the direction taken by some other case-filler in the action.

ex. Head *south* across the desert to Mexico.

SPECIFIC-LOCATION

This is a location associated with the action. Some activities like "ride" take a location which is quite specific to it; there are only certain things which can be ridden in.
ex. Ride *in a dog sled* in Alaska.

INSTRUMENT

The INSTRUMENT is something used as a tool in an action; it is left over afterward.
ex. Cut the butter with *a knife*.

SPECIFIC-RAW-MATERIAL

This is something used and consumed in the course of the action.
ex. Bake a cake with *powdered eggs*.

TOPIC

This case is used for the topic of mental and communication activities.
ex. Talk about *a bear*.

EXCHANGE

This is something received in exchange for another case-filler in the action.
ex. Trade seeds for *food*.

RATE

This case gives a measure of speed, cost, etc. in relation to units or to something else.
ex. Rent a room at *three dollars a day*.

Figure 4.2 (cond.) The major OWL-1 input cases

undertaken. For example, the action conveyed in "Paint the block red," has as principal result that the block is red. The paint brush may also end up red, but this is not the principal result. In Figure 4.1 request-and-respond has its PRINCIPAL-RESULT output case set to the results of the action desired by the AGENT. When the method is successfully completed, the action has occurred and its results obtain.

This takes care of the header and argument specifications; we turn now to the procedural body, the first component of which is the prerequisite. The request-and-respond example has no prerequisite. In general, none have been used in the representation of core dialogue methods. Nevertheless, prerequisites are a useful structure for other sorts of actions. A familiar example of a prerequisite is the requirement that an elementary course of study be completed before a more advanced one is undertaken. Note that prerequisites differ from other entry conditions in that, if a prerequisite is not satisfied, a method is not rejected as a possible course of action. Instead, measures are taken to satisfy the prerequisite; the execution of the method then continues.

The second part of the procedural body is the set of procedure steps. These are either calls to other OWL-1 methods used to model subactions, assertions of states, or calls to "black box" procedures. The last are procedures represented in a conventional programming language (e.g., LISP); these are intended for complex numeric or linguistic computations, or for any procedures whose steps are not to be included explicitly in the model. (We will see a use of "black box" representations below.)

Recall that procedure steps were divided into standard path and recovery path

steps.[25] I will say something more about each. First, standard paths are the normally expected sequences of events for an action. They embody the relatively small number of ways that an action can be concluded successfully, i.e. can be concluded such that any associated goals are met. When an English description of an action is given, it is the standard path steps that are typically included. If deviations from the standard path(s) are described at all, it is only the most important ones that are given. This indicates that the steps in the standard paths of an action embody the minimum of necessary information about it. For example, when giving someone directions for getting somewhere one normally describes only the successful routes. In the normal case, one does not discuss all the many ways someone could get back on the route again after having made a wrong turn. Even when a particular mistake is very common or particularly costly, one usually describes how to avoid it (e.g., "Don't follow the signs") rather than how to recover from it.

To decide which steps belong to standard paths, one can ask whether it would normally be necessary to describe a step when describing the action, or whether it would be necessary to describe some step of which it is a substep. (This second case applies to low-level detail left out because it is either obvious or unnecessary.) Another useful criterion is whether the *absence* of a step is seen as an event. (This criterion is similar to the one used by Schegloff and Sacks in [28] to determine what should go in a sequence.) For example, if someone asks a question and does not get an answer, this is generally worthy of note. If one asks a question and the hearer does not ask for clarification, this absence of discussion is

---

25. This is not quite the whole story. A third type of step is the assertion of a failure condition. Failure conditions are represented as disjuncts off standard paths or recovery paths.

not, in most cases, considered remarkable. The standard path steps, then, are the normally expected sequences of events for an action, those whose non-occurrence is normally considered an event.

It is fine to follow standard path steps as long as the action goes as intended and no expectations are violated. In practice this will probably not be long, and a failure, with its attendant recovery actions, will result.[26] Some recovery actions are very general and should be modelled by independent methods. One type of general example is the recovery from a failure in the basic conditions of a dialogue, e.g., noise on the line in a telephone conversation. Many other failures, however, are quite specific to the task being carried out. Recovery from such failures is modelled by recovery paths in relevant methods. Recovery paths handle domain specific types of failure recovery, and one of their advantages is that they allow the modeller to put in special instructions for doing what is necessary to return to an appropriate standard path.

The standard path steps for the request-and-respond method are steps 1 through 6 in Figure 4.1. Two recovery paths are also shown for request-and-respond, although many more exist to be modelled. The first recovery path in Figure 4.1 could be used to account for a response such as 4.1, and the second could account for the exchange in 4.2 responding to a REQUEST to write a computer program.

4.1 That's not my department. Try Mr. Jones down the hall.

4.2 CO-AGENT: I can only write blocks world programs.
AGENT: That's what I want.

---

26. For my purposes here, I define *failure* as the violation of an expectation. This is slightly broader than its normal usage, and I do not intend for the usual negative connotations to accompany it.

The failures that trigger both of these recovery paths occur in the process of checking the preconditions, step 3, but the recovery paths are associated with the request-and-respond method. Different recovery paths for these failures could be used when a precondition-checking method is called from another method, e.g., request-and-help. Note that the second recovery path itself has an associated failure condition, which would justify the inclusion of an additional recovery path in the appropriate method. Although only the recovery paths in Figure 4.1 contain conditional steps, note that such conditionals can occur in standard paths as well.

One final point that Figure 4.1 illustrates about dialogue methods is their abbreviated form. For any given speech act, only production is represented explicitly, and the understanding process carried out by the other partner is left implicit. For example, the first step of request-and-respond is that the AGENT makes the REQUEST, but no explicit step is given to represent the process of the CO-AGENT understanding the REQUEST. Thus, although the methods are speaker independent in the technical sense defined, they are not without a bias: it does not matter who is specified as the AGENT of a speech act step, but, whoever is, the "story" is told from his point of view. "Listening" steps are left implicit not because they are unimportant, but because the form and timing are predictable. Where a joint model of communication is necessary (e.g., when misunderstandings occur) it is necessary to expand the abbreviated model expressed in the dialogue methods.

This takes care of the request-and-respond example and with it the general description of methods. A few words remain to be said to link methods to the treatment of speech acts described in Section 2. First, note that speech acts such as REQUEST have

their own associated methods. This is in accord with the major thrust of speech act theory, which views utterances not merely as saying but also as doing. The speech act methods have semantic case specifications for an OBJECT, AGENT, and DESTINATION.[27] Here, the OBJECT specification generally corresponds to the propositional content condition identified by Searle. For all speech act methods, the procedure steps are represented by a call to a "black box" procedure for generating utterances. The preconditions of speech acts are represented in the methods as entry conditions associated with the AGENT of the speech act. Preconditions also appear, possibly somewhat restricted, as constraints on the AGENT of the higher level core methods built around the speech acts.

Turning to speech act taxonomy, recall that I adopted the principal intended effect of the speech act as a major taxonomizing principle. This notion of intended effect can be related to core methods via the OWL-I output case PRINCIPAL-RESULT. More specifically, the principal intended effect of a speech act corresponds to the action directly precipitating the PRINCIPAL-RESULT(s) of a core method built around the speech act.[28] For the request-and-respond example, the principal intended effect is step 5, whose results become the PRINCIPAL-RESULT of the core method. [29] We can say then, that core

---

27. DESTINATION is used rather than CO-AGENT because the second participant in a speech act is seen as passive. This does not mean that this participant does nothing, since he or she is doing the work of trying to understand the utterance. By *passive* I merely mean that the second participant is basically receiving the information, rather than playing a major role in shaping it.

28. Where a speech act has more than one associated core method, the principal intended effect will be a generalization that covers the different actions that directly precipitate the PRINCIPAL-RESULT(s).

29. Note that principal intended effects are often, but not always, the last step in a core method. For example, we can construct a core method called state-and-acknowledge,

methods accomodate the taxonomizing principle used for speech acts in Section 2.

***

In summary, methods provide a general representation for actions. In this section, I have emphasized the way that speech acts can be integrated into larger patterns of action, using methods as a representation. In Section 7, method structure will be used in formulating an important class of rules for ISAs.

---

consisting basically of a speech act of stating (done by P1) followed by an understanding and acceptance operation (done by P2). This latter step would be the principal intended effect of the speech act, but it is not customarily the last step in the core method. Instead, the last step of state-and-acknowledge is typically an acknowledgment (done by P2).

5. Three Approaches to the Treatment of Indirect Speech Acts

All of the foundation is in place, and we can now turn our attention to ISAs. To start, I briefly outline three approaches to the treatment of ISAs that have been prominent in the linguistics literature. Too much has been written on this topic for me to attempt a comprehensive summary here, and I intend only to use these approaches to motivate the analysis in Sections 6 and 7.

5.1. Gordon and Lakoff

The first approach of interest here is that taken by Gordon and Lakoff [10]. Concentrating primarily on requests, Gordon and Lakoff propose a set of four sincerity conditions[30] and then give a single powerful rule to account for the different ways that a request can be framed. They say that to make a sincere request, a speaker must want the action done and believe that the hearer can do the action, that he wants to do the action, and that he would not do it if he were not asked to. The first of these sincerity conditions is called *speaker-based* and the next three are called *hearer-based*. The rule given is:

> One can convey a request by (a) asserting a speaker-based sincerity condition or (b) questioning a hearer-based sincerity condition.

This formulation is very attractive because it is so elegant and simple, but it has encountered extensive criticism. Sadock [24] points out that Gordon and Lakoff's rule is

30. It is not clear whether Gordon and Lakoff mean something different by *sincerity condition* than Searle does. At any rate, their choice of sincerity conditions for requesting differs from Searle's.

PRECEDING PAGE BLANK-NOT FILMED

too powerful, admitting utterances that are not requests. For example, 5.1 is not a request for the hearer to move over, even though the similar form 5.2 is.

5.1 Tell me if you can move over.
5.2 Can you move over?

Even if the set of sincerity conditions could be recast to generate only the correct forms, there are still facts that cannot be accounted for by an approach that uses only speech act conditions plus a general rule. Sadock gives the following examples:

5.3 Can you close the door?
5.4 Are you able to close the door?
5.5 Can you please close the door?
5.6 *Are you able to please close the door?

Sentences 5.3 and 5.4 can both be used as requests, and in this usage they are paraphrases of each other. Taking Gordon and Lakoff's approach, both 5.3 and 5.4 would therefore be derived from the same sincerity condition. Sentence 5.6, however, is ungrammatical, and any scheme that derives 5.3 and 5.4 uniformly will be hard-pressed to account for the difference between 5.5 and 5.6.

5.2. Sadock

To answer the objections that he raises against Gordon and Lakoff's theory, Sadock in [24] distinguishes between utterances that have a speech act as their meaning and utterances that mean one speech act but entail another. Thus, according to Sadock, 5.3 is a sort of idiom which entails the notion of request as part of its meaning, while 5.4 is a question according to its meaning, only secondarily entailing a request. Sadock's claim is

that, although rules based on sincerity conditions have their place, they cannot account for the behavior found in speech act idioms.

Sadock's observations seem to be well taken, although his distinction between meaning and entailment appears to me to be ultimately misleading. The difficulty, I think, is that the notion of what a meaning is and what information it should contain are not as well understood as the theory implies. Sadock assumes that each sentence has only one semantic representation that constitutes its meaning, and he himself notes that this makes him unable to account for the deliberate use of ambiguity in conversation.

There is other evidence that Sadock's treatment falls short of a precise characterization of ISAs. Note that some sentences treated by Sadock as speech act idioms have conversational properties based on their surface forms. Consider examples 5.7 and 5.8 used as requests.

5.7 Move over.
5.8 Can you move over?
5.9 O.K.
5.10 I don't know if I can.
5.11 Yes.
5.12 I don't know.

Examples 5.9 through 5.12 are possible responses to 5.8, but only 5.9 and 5.10 are appropriate as responses to 5.7. The indirect request form 5.8, then, can be treated as either a request or a question by the hearer for the purpose of response. Note, however, that 5.8 can still be understood as a request even though it is answered as a question. One may respond with "Yes" and then go on to comply with the request. In fact, to answer "Yes" and then fail to comply indicates either deliberate rudeness or failure to understand the request force of the

utterance. If we follow Sadock and treat 5.8 as an idiom when used as a request, then only the request force, and not the question force, will appear in its semantic representation. It is then impossible to account for the response behavior described.

Another phenomenon that Sadock's theory does not account for is the variation in the indirect forms that are possible for "entailed" requests. For example, Sadock considers questions of the "Are you able to ..." form to entail requests. He also observes that questions of the form "Do you know <interrogative clause>?" can have the illocutionary force (and hence in his scheme the meaning) of the interrogative clause. That is, 5.13 is often equivalent to 5.14 rather than to a question about the knowledge of the hearer.

5.13 Do you know what time it is?
5.14 What time is it?

Putting these two facts together, Sadock's theory would predict that 5.15 should entail a request, but it does not.

5.15 Do you know whether you are able to close the door?

Hence, even forces that Sadock considers entailments of utterances -- and therefore totally meaning-dependent -- seem to have some dependence on surface form.

5.3. Searle

A third approach that is of interest to us here is that of Searle in [30]. I start with a discussion of some generalizations that he proposes and then compare his approach to that of Sadock.

Searle presents four generalizations for directives and five others for commissives. These play a role in Searle's theory analogous to Gordon and Lakoff's single rule, but Searle's generalizations differ from theirs in the following important ways:

1. Searle's generalizations are differentiated according to the parts of the speech act, i.e. propositional content conditions, sincerity conditions, and preparatory conditions. Gordon and Lakoff's sincerity conditions, in contrast, seem to be an amalgam of Searles' sincerity and preparatory conditions.

2. Searle explains an asymmetry in his generalizations by identifying the conversational participant who has the better knowledge of a condition, but he does not state this explicitly in the generalizations. For Gordon and Lakoff, the question of which participant has the better knowledge is at the core of their single rule.

Searle's approach seems to me to be a valuable one, in that the generalizations are more finely differentiated than Gordon and Lakoff's rule. At the same time, I think that Searle's generalizations can be questioned on the count that they are too specific. Recall that the notions of speaker- and hearer-based conditions have not been given an explicit place in the generalizations. Because of this, Searle is forced to state the generalizations in terms of specific types of preparatory conditions, for example, rather than in terms of preparatory conditions in general. Although Searle is correct in looking for finer differentiation than that afforded by Gordon and Lakoff's rule, it appears that his generalizations are needlessly complex. See Section 6 for an alternative proposal.

Searle's generalizations are presented as part of a larger framework. It is instructive to compare this framework with that developed by Sadock, since there are some sharp contrasts. Searle makes the distinction between *meaning* and *use*, saying that an utterance such as 5.16 should be seen as *meaning* a question while, at the same time, it may be *used* as a request.

5.16 Can you pass the salt?
5.17 Are you able to pass the salt?

The fact that 5.16 takes *please* while 5.17 does not is accounted for by a maxim that one should speak idiomatically in the normal case. Thus, Searle sees a request such as 5.16 as "*idiomatic but not an idiom.*" One attraction of this approach is that it accounts naturally (as Searle points out) for the fact that responses to speech acts may correspond to surface form as well as to a conveyed speech act.

There seem to be two major difficulties with Searle's approach. First, although he acknowledges special grammatical behavior of some ISA forms, Searle is not willing to say that such forms "mean" any speech act other than the one corresponding to the surface form. Utterances such as example 5.5 ("Can you please..." forms) are treated as special conventionalized forms, which is to say as *very* idiomatic (but still not idioms). If the *please* in 5.5 is to be accounted for, this apparently must be done in terms of use, not meaning. Thus, we have both use and meaning reflected in some surface forms, which blurs the use/meaning distinction to some extent. To justify such a step, I think we need a more developed account of the notion of "use" and a delineation of its impact on surface form.

Another difficulty with a strict meaning/use distinction appears in the variations possible for a class of ISA forms that will be discussed in detail in Section 7.5. When 5.19 is used in place of 5.18, I assume that Searle would call the statement that is conveyed the meaning and the request that is conveyed the use. How, then, does 5.20 fit into this framework?

5.18 Close the door.
5.19 It's cold in here.
5.20 Must I tell you that it's cold in here?

As I analyze it, example 5.20 can be used to convey three speech acts, first the "Must I tell you...?" ASK, second the "It's cold in here" representative, and third the "Close the door" REQUEST. Note that either response below would be acceptable:

5.21 Oh. I'll close the door.
5.22 All right. I'll close the door.

The following response would also be possible, although rude:

5.23 Yes, you must.

It is not clear how a strict meaning/use approach would account for the three levels of representation involved here. Note that any change to the theory to accomodate "triple message" ISAs will not be simple to make, due to restrictions on degree of indirection for other ISA forms, as exemplified by 5.1 and 5.2 above.

***

It appears, then, that none of the ISA theories presented is totally adequate. All, however, have extremely attractive aspects that are worth preserving. Gordon and Lakoff's approach attempts to derive the ISA forms from general speech act conditions that have independent uses. Sadock, while endorsing such general derivations, emphasizes the special grammatical properties of some ISA forms. Although Searle's overall approach differs from that of Gordon and Lakoff, his "generalizations" can be viewed as a more finely differentiated version of Gordon and Lakoff's (too) simple principle for deriving ISAs. In the approach described in the next sections I have preserved these positive aspects while, hopefully, avoiding the negative ones.

6. Some General Rules for Indirect Speech Acts

One common property of the three theories discussed in the last section is the use of general rules to account for ISAs. While the theories differ in the nature of individual rules and the relative importance given to them, nevertheless general rules for ISAs are included in each case. In light of the large number of ISA forms, such general rules appear to be necessary. I follow Sadock, however, in questioning whether they are sufficient. In this section I present a set of general rules for deriving ISAs from properties of speech acts. These rules are intended as a basis only, and they will need to be both augmented and restricted. Accordingly, in Section 7 I present another set of rules and categories that, combined with the general rules from this section, give a more complete picture of the range of ISAs possible. The reader is asked to keep in mind, then, that this section is intended as only part of a theory, to be amended in Section 7.

The ISA rules presented in this section will be divided into three categories. The categories are based on a common sense view of rational human behavior. The categories are based on the following three maxims of action:

1. One should (only) initiate actions that are necessary.

2. One should (only) initiate actions for which some desirable result or results can be expected.

3. One should only initiate actions that one expects to be possible.

These three maxims will be referred to as the maxims of *Necessity*, *Desirability*, and *Possibility*.

PRECEDING PAGE BLANK-NOT FILMED

Readers familiar with the classic work of Grice on conversational implicature [11] will recognize the approach that is being taken. There, Grice suggests four categories of maxims that are applicable to linguistic actions but which have analogues in other types of actions. Here, I am stating maxims applicable to actions in general but which apply to speech acts as a special case. The maxim of Necessity above has a counterpart in Grice's category of Quantity. The other two maxims have no direct counterparts, and they suggest extensions to Grice's framework.

We will use the maxims of Necessity, Desirability, and Possibility to provide a conceptual organization for ISA rules. Accordingly, the next three subsections discuss rules of each type. Section 6.4 looks at some general behavior with respect to mood and tense. Finally, in Section 6.5 I discuss a fourth type of ISA that may call for a somewhat different approach.

6.1. ISA Forms Related to the Maxim of Necessity

The maxim of Necessity says that one should act when necessary, but one should avoid extraneous actions. When this second clause is applied to speech acts, it yields the following rule:

Rule 6.1

PI can convey a speech act indirectly by --

(i) ASKing whether the intended speech act is necessary

e.g., the REQUEST "Do I need to tell you to shut the door?"

(ii) ASKing whether an equivalent speech act (i.e., one with the same principal intended effect) has already been performed

e.g., the REQUEST "Did anyone ask you to take out the garbage?"

(iii) ASKing whether the principal intended effect can be expected to occur without the speech act

e.g., the REQUESTs:
"Are you planning to take out the garbage?"
"Are you going to take out the garbage?"

(iv) ASKing whether the principal intended effect of the speech act has already occurred

e.g., the REQUESTs:
"Did you take out the garbage?"
"Have you taken out the garbage?"

For Rule 6.1 and other rules presented in this paper, I am assuming a relatively constrained relationship between the actual representations of the rules and the semantic representations of individual utterances. See Appendix B for a description of this relationship. Although REQUEST examples were used in 6.1 (and will continue to be used in this section), the general ISA rules are intended to apply to any speech act where P2 can be expected to have the appropriate knowledge to respond to P1's ISA.

In Rule 6.1, note that Clause (iv) is also expected to account for forms such as 6.1 used as a REQUEST to take out the garbage:

6.1 Is the garbage out?

This will require additional rules and distinctions, however, which will be introduced in Section 7.5.

Earlier versions of this paper accounted for the "Will you <action>?" REQUEST form by an appeal to Clause (iii) of Rule 6.1. I have not been satisfied with this account, however, because some uses of the form are not motivated by questions of the necessity of the action. Consider, for example, 6.2.

6.2 Will you accept a ride to the airport?

This example can be accounted for by the maxim of Possibility. One can view 6.2 as P1 asking P2 whether the outcome of an OFFER by P1 will be successful (i.e., acceptance). While some readings of the "will" form may be accounted for by Rule 6.1, we need a broader view of this form. See Section 6.3.

Finally, note that P1 is permitted to use an ISA only when P1 can reasonably expect P2 to decifer P1's intent, i.e., to recognize the indirection. Neither Rule 6.1 nor any of the other rules presented here, however, includes this information. It appears that this constraint is part of a more general constraint that P1 avoid ambiguity. That is, P1 is obligated -- to the best of his or her ability -- to frame *any* utterance (ISA or not) in such a way that P2 can understand the message that P1 intended to convey. See Grice [11] for discussion of an "avoid ambiguity" maxim.

6.2. ISA Forms Related to the Maxim of Desirability

Next we come to the maxim of Desirability, restated here.

> One should (only) initiate actions for which some desirable result or results can be expected.

Related to this maxim, we have the following general ISA rule:

Rule 6.2:

P1 can convey a speech act indirectly by --

-- a representative with the propositional content that some desirable result or results can be expected for some intended effect of the speech act.

e.g., the REQUEST "I would be happier if you'd substantiate those figures." Here, the desirable result is the happiness of P1, and the intended effect of the REQUEST is that P2 substantiate the figures.

Again, for Rule 6.2, although the example is a REQUEST, the rule is expected to apply to speech acts in general.

In Rule 6.2, note that the intended effect need not be an immediate RESULT or PRINCIPAL-RESULT (see Section 4) of the speech act; it may be several times removed in the causal chain. Similarly, the desirable result need not be an immediate RESULT or PRINCIPAL-RESULT of the intended effect. Thus, we have the following examples of indirect representatives.

6.3 You'll be happy to hear that Claudia won.
(*Hear* refers to a RESULT of the speech act.)

6.4 You'll be happy to learn that Claudia won.
(*Learn* refers to the principal intended effect of the speech act.)

6.5 You'll be happy to know that Claudia won.
(*Know* refers to the PRINCIPAL-RESULT of the principal intended effect of the speech act.)

Note that the desirable results from the maxim need not relate only to P1 and P2. Consider the following REQUEST example, taken from [14].

6.6 This room would look a lot better if you dusted it.

Example 6.6 contains information about an inanimate object with the claim that it would be more attractive (to anyone) if the REQUESTed action were performed. Neither P1 or P2 is explicitly mentioned in the statement of the result.

6.3. ISA Forms Related to the Maxim of Possibility

The third maxim proposed was the maxim of Possibility: one should only initiate actions that one expects to be possible. The minimal conditions for this maxim to obtain can be stated in terms of the framework developed in Section 4. A speech act can be considered to be possible if:

1. the preconditions of its associated method can be satisfied.

2. for an action-centered speech act, the semantic input cases of the action named can be filled.

3. for an action-centered speech act, the prerequisites of the action named can be satisfied.

The maxim of Possibility has a rich set of corresponding ISA rules. Rules related to (1) and (2) above are explored in the next two subsections. Rules related to (3) above are discussed in Section 7.5.2.

6.3.1. Forms Based on Preconditions of Individual Speech Acts

The first set of ISA forms based on the maxim of Possibility are those derived from preconditions of speech acts. Note that the preconditions associated with the speech acts

give conditions *particular to the speech act*. There are other general conditions of speech acts which must obtain for the individual act to be well-formed, as evidenced by rules 6.1 and 6.2. I am not including these when I refer to preconditions. The approach taken in this subsection will be to distinguish three classes of precondition and formulate seven rules using the classes distinguished.

While the rules suggested are intended to apply generally to speech acts, once again REQUEST will be used to supply examples. (Examples for OFFER, SUGGEST, and ASK are given in Appendix C.) To start, then, recall the preconditions identified in Section 3 for REQUESTs:

I. P1 wants P2 to take responsibility for carrying out the action.

II. P1 believes that P2 can take responsibility for carrying out the action.

III. P1 believes that P2 is willing to take responsibility for carrying out the action.

IV. P1 believes that P2 is obligated to P1 (and possibly to others) to take responsibility for carrying out the action. (This obligation may be a role obligation, an authority obligation, or a general obligation to be helpful.)

The preconditions of REQUEST, and other speech act preconditions, can be divided according to *which dialogue participant has the better knowledge of the conditions specified.*

I. P1-based preconditions.

Here P1 has inherently better knowledge of whether or not the topic of the precondition holds. The *topic* of preconditions that begin here with "P1 believes" is considered to be the direct object of the initial "believe." For other preconditions (e.g., REQUEST I.), the topic is the entire pattern. Preconditions that are P1-based include those whose topic is a goal of P1 (e.g., REQUEST I.).

2. P2-based preconditions.

Here P2 has inherently better knowledge of whether or not the topic of the precondition holds. Preconditions that fit this category include P1's beliefs about P2's intentional states. An example of a P2-based precondition is REQUEST (III).

3. Unmarked preconditions.

For these preconditions, determination about which participant has the better knowledge of the precondition depends on properties of the particular speech act and/or its context. Examples are REQUEST (II) and (IV).

Using these precondition types, we can construct the following seven rules for ISA forms.

P1 can convey a speech act indirectly by --

Rule 6.3 (for P1-based preconditions):

-- a representative of the topic of a P1-based precondition of the speech act.[31]

e.g., I want you to water the plants. (REQUEST I.)
I hope you will use common sense. (REQUEST I.)

Rule 6.4 (for P2-based preconditions):

-- an ASK of the topic of a P2-based precondition of the speech act.

e.g., Do you want to shut the door? (REQUEST III.)

Rule 6.5 (for unmarked preconditions):

-- an ASK of the topic of an unmarked precondition of the speech act.

---

31. It is probably the case that for all P1-based preconditions, the topic is equivalent to the entire precondition. This rule could probably be stated simply as "a representative of a P1-based precondition," but for now I stay with the more general version.

This rule applies in a context where P1 believes P2 has better knowledge of the condition in the precondition topic.

e.g., Is it your turn to do the dishes? (REQUEST IV.)

Rule 6.6 (for unmarked preconditions):

-- a representative of the topic of an unmarked precondition of the speech act.

This rule applies in a context where P1 believes P1 has better knowledge of the condition in the precondition topic.

e.g., It's your turn to do the dishes. (REQUEST IV.)

Rule 6.7 (for unmarked preconditions):[32]

-- a representative of an unmarked precondition of the speech act where P1 believes he or she has better knowledge of the condition.

e.g., I believe it's your turn to do the dishes. (REQUEST IV.)
I assume you are able to shut the door. (REQUEST II.)

Rule 6.8 (for groups of preconditions):

-- a REQUEST form of an action that is a goal of P1 (i.e., A for "P1 wants A").

This rule is applicable only when the speech act has preconditions that are exact matches or specializations of the four preconditions of REQUEST.

e.g., Take a cookie. (OFFER IV.-VII)

---

32. Utterances derived from this rule could also be treated as double indirections, e.g., an indirect form of a representative of the speech act. If this is done, however, these utterances would be special cases, since double indirections are not possible in general. (Recall example 5.1, "Tell me if you can move over," which is not a REQUEST that P2 move.) Whatever the treatment of these utterances, note that the initial "I believe," "I assume," etc. functions more as a mitigator than as a separate speech act. For a discussion of mitigators, see [14].

Rule 6.9

-- an ASK about whether P2 will take responsibility for carrying out an "active" action that is a goal of P1 (i.e., A for "P1 wants A"). [33]

This rule is applicable only when the speech act has preconditions that are exact matches or specializations of the four preconditions of REQUEST.

e.g., Will you accept a ride to the airport? (OFFER IV.-VII.)

Although these seven rules are more complex than Gordon and Lakoff's single rule, they give us a more powerful basis for identifying the set of precondition-based ISA forms. First, nothing corresponding to Rules 6.8 and 6.9 (which operate on groups of preconditions) is found in Gordon and Lakoff's theory. Second, the treatment of the precondition classes is more powerful here. Nothing in Gordon and Lakoff's dichotomy between speaker- and hearer-based sincerity conditions predicts conditions that can either be questioned or asserted, depending on context.

The major difference between Rules 6.3 to 6.9 and Searle's generalizations rests in the level of generality of the two schemes. Recall that Searle's generalizations are more numerous because they are associated with different classes of speech acts. In contrast, the rules given here apply to speech acts in general, and they are proposed as the full set of general ISA rules that apply to preconditions.

In terms of specific rules, rules 6.8 and 6.9 differ most from Searle's generalizations, because, again, they operate on groups of preconditions. Rule 6.9 is of special interest. In

---

33. By "active" I mean an action that is viewed as an event in the environment shared by P1 and P2. For further discussion, see Section 6.4.

Searle's scheme, the very common "Will you <action>?" form is derived from the propositional content condition of directives. I feel that this approach lacks semantic motivation, and so the approach taken here is to appeal to the maxim of Possibility. One can interpret a "will" question as a question about how P2 will respond to P1's speech act. The "will" ISA form can be seen as a hedge in case P2 cannot respond appropriately, does not wish to do so, or recognizes no obligation to do so. By using this form, P1 can carry out the intended speech act even though he or she may have some reservations about its effects. Anchoring "will" forms in groups of preconditions and appealling to the maxim of Possibility provides these forms with a stronger semantic motivation than the more structural account offered by basing rules on propositional content conditions.

Given the constraints on matching described in Appendix B, Rules 6.3 through 6.9 are relatively restrictive. I will conclude this subsection with some observations on the forms that the rules *do not* account for. First, the rules as written do not account for differences in tense and mood. That is, Rule 6.5 generates example 6.7 but not 6.8 and 6.9.

6.7 Are you able to drive Sarah to school?
6.8 Will you be able to drive Sarah to school?
6.9 Would you be able to drive Sarah to school?

Since 6.8 and 6.9 are legitimate indirect REQUESTs, they must be accounted for; Section 6.4 contains proposals for handling this tense and mood behavior.

Another collection of forms that is not covered by Rules 6.3 to 6.9 are those that contain negations of the sort found in patterns 6.10 to 6.13.

6.10 Won't you <action>?
6.11 Couldn't you <action>?
6.12 Shouldn't you <action>?
6.13 Can't you <action>?

Application of Rules 6.3 to 6.9 to the REQUEST preconditions does not yield such negated forms. In this case, I think that the exclusion is well-founded. Forms 6.10 to 6.13 can be interpreted as conveying several different speech acts, but two of the more prominent ones are begging and persuading. What these two speech acts have in common is PI's need to overcome some sort of resistance in P2. (The resistance is to the action named in the speech acts; both are action-centered.) REQUEST does not involve overcoming any special resistance; precondition (III) relates to PI's estimate of P2's openness to the action. Thus, while forms 6.10 to 6.13 are clearly similar to REQUEST forms, they are not treated here as REQUESTs.

6.3.2. ISA Forms Based on Semantic Cases of Actions

We continuing with the maxim of Possibility and observe that ISA forms for some types of action-centered speech act are based on semantic input cases of the action named. Consider, for example:

6.14 Do you have *a rope* to throw to that man?
(OBJECT of the throwing action)

6.15 Have you got *a hammer* to fix that picture?
(INSTRUMENT of the fixing action)

6.16 Do you have *enough gas in your car* to drive George to the airport?
(SPECIFIC-RAW-MATERIAL of the driving action)

All of these examples can (among other things) act as indirect REQUESTs for P2 to do the action named. Given the appropriate context, they can also act as indirect SUGGESTs. (One such context would be the case where P2 is trying to decide what to do in a situation.)

The basic pattern that these examples fit is given in the following rule:[34]

Rule 6.10

P1 can convey a REQUEST or SUGGEST indirectly by --
-- ASKing whether P2 has <item that fills a semantic case slot> for/to <action>

Here, the semantic case is one defined for the action named. It is not clear whether this formulation is too broad; a pattern involving only a subset of the semantic cases may be more correct. For instance, *a partner* in 6.17 is a filler for the CO-AGENT semantic case. Example 6.17 uttered by a tournament official can definitely be a REQUEST that P2 get a partner, but can it also be used as a REQUEST that P2 play the next match (as Rule 6.10 would have it)?

6.17 Have you got a partner to play the next match?

While example 6.17 suggests that Rule 6.10 is too broad, other evidence suggests that it is too restrictive. In addition to examples that correspond quite well to the semantic input cases identified in Section 4, there are other examples that are clearly related but not as easy to account for. Consider:

6.18 Do you have time to take Sarah to the airport?
6.19 Do you have room to store my plants?

Examples 6.18 and 6.19 are time- and space-related, respectively. When representing actions with OWL-1 methods, it is assumed that all actions occur at some time, but, because of the

34. It is possible that this rule can be generalized, since analogous, but not always identical, forms exist for other action-centered speech acts.

generality of this attribute, time is not represented as a semantic input case. (This does not preclude the use of a different representation to associate times with particular events.) Even if the starting time of an action were represented as a semantic input case, this would not be an immediate solution to the problem of accounting for example 6.18. This is because the time discussed there is *elapsed time*, not a single point in time. Similarly, example 6.19 can be related to the semantic case SPECIFIC-LOCATION, but it seems to be a question about the *size* of a place, rather than simply a place. While we could consider extending Rule 6.10 to include attributes of semantic cases, we run the risk of admitting other forms that are not ISAs. Pending a solution to the difficulties posed by examples 6.17 to 6.19, I leave Rule 6.10 as stated. This is done, however, with the caveat that it merits further investigation.

To sum up this subsection, we can say that for at least some action-centered speech acts, ISA forms are based on semantic input cases of the action named. Work remains to be done to establish the exact subclass of action-centered speech acts that permit this type of ISA form.

6.4. Tense and Mood Variations

This completes the presentation of ISA rules associated with the three maxims of action. In this subsection I briefly discuss the derivation of future tense ISAs, then go into somewhat more detail for subjunctives.

The following rule expansion can be applied to all ISA rules:

Expansion for future tense:

If an action-centered speech act refers to an action that is to take place sometime

other than "right away," then the ISA form may be stated in the future tense, if this is appropriate.

Although the future expansion and the subjunctive expansion below are rules, note that they are not at the same level as the other ISA rules presented so far. The expansions are rules that act on other rules. Alternatively, it would be possible (but wordy) to write individual ISA rules with separate clauses to account for tense and mood behavior.

In the future expansion, the "if this is appropriate" hedge is used to account for differences such as those between the REQUEST forms 6.20 through 6.23.

6.20 Will you be able to <action>?
6.21 Can you <action>?
6.22 Will you be willing to <action>?
6.23 ?Will you want to <action>?

While 6.20 and 6.22 occur in the future tense, there is no future for *can* and 6.23 is questionable as an indirect REQUEST. Some properties, apparently, are viewed as more time-variant others. It is this difference that the "if appropriate" hedge is meant to allow for.

Turning to the subjunctive, we can frame the following rule expansion:

Expansion for subjunctive:

ISA rules that generate ASK forms where the state or action that will be named by the finite verb is not set in the past may also be used to generate subjunctive forms if the conveyed speech act has a principal intended effect that involves active response on the part of P2.

By "active" I mean a response that is viewed as an event in the environment shared by PI

and P2. This constraint is used to include utterances such as the indirect ASK 6.24 while excluding those such as the indirect representative 6.25.

6.24 Would you know if I can get a boarding pass at that desk?
6.25 *Would you know that the plane left already?

I am assuming that both of these utterances are related to "Do you know..." forms. Here, the principal intended effect of 6.24 is the "active" response of supplying an answer, while the principal intended effect of 6.25 is the more passive response of understanding the propositional content.

The subjunctive expansion above is written to exclude past actions. This is relevant for forms generated from Rule 6.1. For example, although form 6.26 is a common type of indirect REQUEST, 6.27 is not an indirect REQUEST at all.

6.26 Have you dusted?
6.27 Would you have dusted?

The exclusion of past actions in the subjunctive expansion effectively bars the interpretation of 6.27 as a REQUEST to dust.

An important question that is relevant here is what event or condition is predicated by speech act-related subjunctives. This question has received some attention in the literature One common answer to this question is that the implicit event is "if PI performed the speech act", e.g., "if I asked you" for REQUEST. Searle points out in [30], however, that while 6.28 may convey a request, 6.29 may not.

6.28 Would you pass me the salt?
6.29 Would you pass me the salt if I asked you to?

Searle suggests instead that the subjunctive is predicated on "if you please" or some variant. The notion of active response used in the subjunctive expansion above complements this interpretation. "If you please" is a way for P1 to ask P2 to look favorably on an action, and that action is the active response associated with the successful completion of the speech act.

6.5. Examples that Fall Outside the Rules

The final topic for this section is a group of ISAs that are not completely accounted for by the approach that has been taken. Consider, for example:

6.30 Is your leg healed enough for you to go to the store for me?
6.31 Will you be home in time to walk the dog?

Looking back at the REQUEST preconditions, these examples seem to correspond to (II), related to P2's capability, but the correspondence cannot be accounted for by the rather restricted matching relationship I am assuming (see Appendix B).

In trying to account for examples such as these, my first inclination was to appeal to the definition of *can* in Precondition (II) and expand the matching rules to include matches on parts of definitions. That is to say, the semantic representation for *can* would be assumed to have a definition that included notions of physical ability, having free time, and being within an appropriate spatial range to do an action. It turns out, however, that this treatment puts a great deal of strain on the precondition-based approach without really solving the problem exemplified by 6.30 and 6.31. In terms of matching, there is still a good distance between a healed leg and the more general notion of physical ability, or between the relationship of being at home with a dog and the idea of being in the right range to

perform the action of taking it for a walk. There are definite correspondences here, of course, but the point is that establishing the correspondences can be a complex computational problem. This is in contrast to the approach I have been taking, where the matching process has been carefully constrained.

Although the approach proposed here appears to be too restricted to handle some cases, the answer does not necessarily lie in a wholesale replacement of Rules 6.1 to 6.10 and Appendix B with a more general formulation. Instead, I advocate a layered approach, using several sets of rules of varying power. We will have made some progress if we can isolate the classes of ISAs that can be derived from properties of speech acts using highly restricted rules. These forms constitute a significant class of ISAs, and their restricted nature can be exploited in a computational model. Other cases can then be handled by a more general, and presumably more expensive, inferential mechanism. An example of a general inferential approach is found in Allen [1].

7. A Taxonomy of Indirect Speech Act Forms

7.1. Overview

In this section I present a taxonomy of speech act forms, with emphasis on indirect speech acts. The sets of categories in the taxonomy are to be viewed as different dimensions, so that utterances can, and typically do, belong to more than one category at once. I will use REQUEST as a source of examples. The categorization scheme does apply to other speech acts, often, however, with some simplification necessary.

The first distinction that will be discussed is whether the form is *direct*, *indirect*, or *tag*. Indirect forms, of course, are of primary interest for this paper. It is also, however, necessary to consider direct speech act forms for the same reason that it was necessary to clarify speech act organization in Section 2: one speech act's direct form is another speech act's indirect form. Besides direct and indirect forms, I distinguish tag forms. Tag forms are sometimes considered indirect forms, but I will argue for their treatment as a separate class.

The next distinction that will be considered is whether an utterance is *single-force*, *double-force*, or *triple-force*. Corresponding to this is the notion of whether a given speech act is conveyed as the *immediate*, *secondary*, or *tertiary* force of the utterance. This second distinction corresponds roughly to Sadock's meaning/entailment distinction, but it is recast to avoid the problems described in Section 5. Related to the distinction among forces is the question of the existence of speech act idioms. In Section 7.4 I take a position on this issue with the introduction of *frozen ISA forms*.

The final two distinctions, which apply only to what I have called action-centered speech acts, have received only casual attention in the literature. The first distinction hinges on whether PI names the action involved explicitly or refers to it only implicitly. The two ISA categories are, accordingly, *explicit-action* and *implicit-action*. For implicit-action ISA forms, there is the further categorization into three types, which for simplicity will be referred to with numbers. The difference between type 1, 2 and 3 implicit-action forms lies in the type of information needed to recognize the action intended by PI.

We can now proceed to a more detailed discussion of each set of categories.

7.2. Direct Indirect and Tag Forms

Figure 7.1 illustrates the first distinction for REQUEST forms. In this figure, slashes are used to indicate choices, parentheses indicate optional words or phrases, and angle brackets contain descriptions of words or phrases.

Direct forms come in two varieties: *performative* and *simple*. The former is Austin's class of utterances in which the speech act is stated explicitly. Simple forms are special grammatical or idiomatic structures that are associated with the speech act. The simple form for a REQUEST is the imperative. Note that the relationship between simple forms and speech acts is not one-to-one. The simple form for giving instructions is, like that for REQUESTs, an imperative, although possibly modified by an infinitive phrase (e.g., "To open, push in and twist.") Warnings also use the imperative as a simple form. Not only do some speech acts share simple forms, but other speech acts have more than one, e.g., dialogue openers with the simple forms "Hello," "Hi," etc. Finally, many speech acts, particularly ceremonial ones, lack simple forms altogether and rely on the performative for a

DIRECT FORMS

PERFORMATIVE
I request that <action>.
ex. I request that you write a program to manipulate blocks for me.
SIMPLE FORM
<Imperative of action requested>
ex. Write a program to manipulate blocks for me.

INDIRECT FORMS

(1) Will <action>?
ex. Will you write a program to manipulate blocks for me?

(2) Would <action>?
ex. Would you write a program to manipulate blocks for me?

(3) Can <action>?
ex. Can you write a program to manipulate blocks for me?

(4) Could <action>?
ex. Could you write a program to manipulate blocks for me?

(5) I want <action infinitive>.
ex. I want you to write a program to manipulate blocks for me.

(6) I would like <action infinitive>.
I'd like <action infinitive>.
ex. I would like you to write a program to manipulate blocks for me.

TAG FORMS

(1) <imperative of action requested>, will you?
ex. Write me a program to manipulate blocks, will you?

(2) <imperative of action requested>, would you?
ex. Write me a program to manipulate blocks, would you?

(3) <imperative of action requested>, can you?
ex. Write me a program to manipulate blocks, can you?

(4) <imperative of action requested>, could you?
ex. Write me a program to manipulate blocks, could you?

Figure 7.1. Some examples of explicit-action forms for REQUEST

direct expression of the speech act.

The next class is indirect forms, in which the speech act is conveyed by a simple form for another speech act. Figure 7.1 lists a small subset of the possible indirect forms. The other sets of categories presented in this section will give us a way to look further at the nature of indirect forms. Note that Figure 7.1 does not include negations of these indirect forms, e.g., those involving *can't, won't, couldn't,* etc. Recall the decision discussed in Section 6.3.1 to treate these forms as ISAs for the related speech acts for begging and persuading (among other possibilities).

Finally, tag forms are constructed from a simple form combined with the identifying part of an indirect form. The identifying part for most indirect forms is the finite verb plus subject (e.g., "Move it over, *would you*?"), but in some cases it encompasses more (e.g., "What's the answer, *I'd like to know*."). The ISA forms that may have corresponding tag forms can be characterized more specifically as *frozen forms*, a class which is defined in the next subsection. For REQUESTs, the tag forms correspond to the interrogative frozen ISA forms. This is not, however, necessarily the case for other speech acts. Since I can give no more precise specification of the ISA forms that have corresponding tag forms, this is left as a problem for further research.

A separate category has been allotted for tags, since they may function as an amalgam of two forms. (For example, in many contexts "Move over, would you?" has the brusqueness of the simple form which is then softened somewhat by the last-minute introduction of the more humble "would you".) A view which treats tag forms merely as transformations of indirect forms misses the dual nature of the message that they can convey.

7.3. The Speech Acts Conveyed

The next set of categories gives a way to talk about the speech act or acts conveyed by an utterance. My first assumption is that some, but not all, utterances convey more than one speech act. These different speech acts will be referred to as the *forces* of an utterance. For the utterances that convey multiple speech acts, the largest number of such speech acts that I have found in English is three. Utterances are therefore referred to as either single-force, double-force, or triple-force, and the three forces will be called *immediate*, *secondary*, and *tertiary*.[35] The speech acts are called only *forces*, since I wish to avoid appealing to any notion of "meaning." Although if pressed I would have to say that I see all speech act forces conveyed by the utterance as part of the meaning of the utterance, semantic investigations are at too early a stage, I think, for a concrete definition of meaning.[36]

Examples of utterances that convey different numbers of forces are easy to find. First, utterances with only an immediate force include direct, simple forms, e.g., 5.18 (repeated below) uttered with an immediate REQUEST force. An utterance with both immediate and secondary force might be 5.4 uttered with an immediate ASK force corresponding to the surface form and a secondary REQUEST force. Finally, there is example 5.20 which may be uttered with the intent to convey an immediate ASK, a secondary representative, and a tertiary REQUEST force.

5.18 Close the door.
5.4 Are you able to close the door?
5.20 Must I tell you that it's cold in here?

---

35. My use of the term *secondary* differs from its use in Searle [30]. I hope that this will not be a source of confusion.

36. See, however, Zwicky and Sadock [36] and Sadock [25] for important work in this area.

To account for the different forces in these examples, let us hypothesize a mechanism that applies the rules given in Section 6. (I postpone to Section 7.5 the issue of how to link the statement of the condition in 5.20 with the particular action intended.) With this hypothesized mechanism, immediate forces of ISAs are derivable directly from the surface form (or a semantic representation of it), secondary forces are derived by one application of the rules, and tertiary forces are derived by two applications of the rules.

There are two basic difficulties with the mechanism proposed, both of which are familiar from Section 5. First, recall Sadock's observations that 5.1 (repeated below) is not an indirect REQUEST.

5.1 Tell me if you can move over.

We need some way to block the assignment of three forces -- REQUEST of a tell, ASK, and REQUEST of a move -- to 5.1. At the same time, it is necessary to preserve the possible triple-force reading of 5.22. The second difficulty with the proposed mechanism is illustrated by the examples repeated here:

5.5 Can you please close the door?
5.6 *Are you able to please close the door?

With the mechanism proposed, both 5.5 and 5.6 would be produced in the same way, and there is no way to account for their differing behavior with respect to *please*.

To take care of the first problem with the proposed mechanism, we must first amend the rules given in Section 6. Each rule must specify that the speech act form used must be a *simple* one, that is, interrogative for ASK, declarative for the different representatives, and imperative for REQUEST. For example, Rule 6.3 is now:

Rule 6.3 (for PI-based preconditions):

PI can convey a speech act indirectly by --

-- a *simple form* of a representative of the topic of a PI-based precondition of the speech act.

In addition to this change, the proposed mechanism must be limited to a single application of the rules for all the ISA forms considered here except for one subclass. The only case where double indirection, and hence the production of tertiary force, is uniformly possible is a class that will be called *type 3 implicit-action forms.* [37] This class of ISA forms will be discussed in Section 7.5.3, and in 7.5.4 consideration is given to the question of why such double indirection should be possible for this class.

This brings us to the second difficulty with the proposed mechanism. The special behavior illustrated by 5.5 and 5.6 is one example of the type of phenomenon that the idea of speech act idioms attempts to account for. Although a relatively small group of ISA forms display such special behavior, they have received a great deal of attention in the linguistics literature. Because of a rather interesting combination of characteristics, these forms have been resistant to attempts to fit them into various theoretical frameworks. After several different iterations I find that I come close to subscribing to the concept of speech act idioms. I feel, however, that Searle has made a convincing argument against the use of

37. As ever, there are some borderline cases. Some indirections on secondary forms do seem to carry three forces, e.g., "I'd like to know if you could drive us to the airport" (an immediate STATE, a secondary ASK, and a tertiary REQUEST). The examples I have found tend to be isolated, however, and they will be treated here as frozen forms (see the next subsection). The only uniformly productive class among those considered in this paper are the type 3 implicit-action forms.

the term *idiom*. The next subsection, therefore, describes a slightly different class called *frozen ISA forms*.

7.4. Frozen ISA Forms

I start the explication of frozen ISA forms with the intuition that the sentence in 7.1 can be uttered as a REQUEST -- not an ASK and a REQUEST, but simply a REQUEST.

7.1 Will you close the door?

The first time that I consciously thought about example 7.1, it took a fair amount of effort for me to get to the literal speech act force (a question about a future action). Not all ISAs are like this of course. An utterance such as 5.19, repeated below, seems to be primarily a representative, and only secondarily a REQUEST.

5.19 It's cold in here.

In searching for the source of this intuition, I come to the fact that the ISA forms that I want to call "frozen" can or do display special behavior uncharacteristic of the rest of the surface form. I therefore define frozen forms as follows:[38]

> A frozen ISA form is an ISA form with an immediate force *that does not correspond to the surface form*. For such an utterance, one or both of the following obtains:

---

38. It would be nice if special intonational behavior were associated with indirect interpretations of utterances, as well. The evidence, however, points to special disambiguating intonations associated with some *direct* interpretations of utterances. (See Sag and Liberman [26].) Although this study reports a strong association between some intonation contours and indirect interpretations, the intonations did not decisively disambiguate indirect forms.

1. It can or does display syntactic behavior atypical of the rest of the form.

2. It can or does display cooccurrence behavior atypical of the rest of the form.

This definition of frozen forms is almost a direct translation of Sadock's criteria for speech act idioms into the framework that has been developed here.[39] (An addition to this formulation, however, is proposed below.) Clause 2 above is illustrated by the special cooccurrence properties of example 5.5 as contrasted with example 5.6 (repeated here).

5.5 Can you please close the door?
5.6 *Are you able to please close the door?

An example pointed out by Sadock that fits Clause 1 is 7.2, which is a passive form of 7.1 above.

7.2 Will the door be (*please) closed by you?

The REQUEST force in 7.1 is, intuitively, more direct than that in 7.2. Moreover, the special cooccurrence properties of 7.1 are lost.

The presence of *please* in 5.5 is not an isolated phenomenon, as Sadock has pointed out. Compare:

7.3 Please, will you close the door?
7.4 Will you please close the door?
7.5 Will you close the door, please?
7.6 Please, it's cold in here.
7.7 *It's cold in here please.

---

39. Note that Sadock's "paraphrase" criterion ([24], chap. 5) has been omitted from the frozen form definition. While the notion of a paraphrase is useful intuitively, I feel that we lack a refined understanding of what does, and does not, constitute a paraphrase of an utterance. The notion of a paraphrase is therefore not yet a computationally useful one.

Even where initial *please* is possible for secondary REQUESTs (e.g., 7.6), Sadock points out that its sense is slightly different. This use of *please* is more likely to appear in a REQUEST that someone stop doing what he or she is either about to do or currently doing. In addition, for many immediate REQUEST forms, *please* may occur initially without a pause, e.g.,

7.8 Please will you close the door?
7.9 *Please it's cold in here.

The non-imperative forms that convey a REQUEST and take *please* freely include the indirect forms listed in Figure 7.1. Note, however, that not all cases are as clear-cut as the examples given. Consider:

7.10 *Would you mind please picking up George at the airport?
7.11 ?Would you mind picking up George at the airport please?

There are a number of these cases for which terminal *please* seems to be at least marginally acceptable. This sort of gradual decay of a criterion, as opposed to a sharp cutoff, is a familiar phenomenon in linguistics. The way that frozen forms have been defined, however, it is not necessary that the frozen forms for a speech act all display a *single* type of special behavior or that, in this case, *please* occur freely. It is enough that the form may display *some* syntactic or cooccurrence behavior atypical of the rest of the form.

Given the definition of frozen ISA forms, we can amend the simple mechanism hypothesized above. For each speech act, it appears that the set of frozen forms must be represented by its extension. (This applies to the *forms*, of course, not to the set of the individual utterances.) The set of patterns for frozen forms can be used to derive

immediate force, given the surface form of the utterance. For frozen forms, this process occurs *instead of* the normal application of the general ISA rules.

This is an improvement, but the proposed mechanism is still not complete. Since frozen forms have an immediate force that does not reflect the surface form, I must account in some way for the response behavior discussed in Section 5.2. It will not do in general to say that this response behavior results from processing frozen forms as if they were not frozen, e.g., as if the REQUEST force for 7.1 were secondary. If this were done, we would be at a loss to account for the special surface behavior such as the unexpected presence of *please* in a non-imperative frozen REQUEST form.

The best answer, I think, lies in assuming that each frozen form pattern has associated with it a pointer to the relevant general ISA rule, particularized to a speech act appropriate to the surface form. The pointer provides a potential interpretation of the utterance, that may or may not be taken. Given this pointer, no information is lost, and responses may be keyed off either the immediate force or the general ISA rule.[40] Such a pointer is needed only for frozen forms. In other cases, immediate force corresponds to surface form, and keying responses off immediate force is equivalent to responding to surface form.

As a final remark on frozen forms, note that they fall somewhere between Sadock's speech act idioms and Searle's idiomatic but non-idiom forms. It is the addition of the

---

40. This treatment of the literal speech act force of a frozen form as only a potential force is comparable to Morgan's notion of *short-circuited implicature* as it is applied to speech acts [21]. Short-circuited implicatures are those whose "literal meaning is 'latent' rather than the basis for an inference." The pointer to an ISA rule proposed here for frozen forms appears to combine information from the occasion and purpose slots in Morgan's scheme.

pointer that distinguishes frozen ISA patterns from patterns of idioms. For these "true" idioms, the literal interpretation of the form is totally lost in normal usage, e.g., "How do you do?" used in an introduction exchange. For frozen forms, on the other hand, the pointer makes a literal interpretation accessible, if not always accessed. While frozen forms are not idioms, they are also not merely idiomatic, because, by virtue of possible "cross-over" syntactic and/or cooccurrence behavior, they are atypical of the speech act that corresponds to the surface form.

7.5. Explicit and Implicit Actions

We turn now to action-centered speech acts and consider the distinction between actions named explicitly by P1 and those referred to only implicitly. Explicit forms are the straightforward case; example 5.18 ("Close the door.") is a typical explicit form. Implicit-action forms are relatively complex; at least three different varieties of implicit-action speech act can be identified. In this subsection, I discuss implicit-action ISAs, taking each of the three types in turn. Due to the complexity of the subject matter, the account will be fairly detailed.

Before plunging in, a word must be said about the relationship of implicit-action ISAs to dialogue context. Implicit-action ISAs rely heavily on context. P1 is obliged to frame an implicit-action ISA in such a way that P2 can uniquely identify the action, given the context. P1 must take into account not only the context, but also P1's knowledge of P2's model of the context. Dialogue context plays a progressively greater role for each of the three implicit-action types that will be described here.

Despite this heavy reliance on context, it appears that the difference between implicit-action and explicit-action utterances in this respect is basically *quantitative* rather than *qualitative*. Although more information must be supplied by context for implicit-action ISAs, the same *types* of knowledge are needed for both classes of utterance. This knowledge is needed both to choose the correct interpretation for an utterance (e.g., to distinguish between direct and indirect speech acts) and to round out the specification of an action. See Section 12 for a discussion of some of the knowledge sources that come into play for all types of ISAs, including both implicit- and explicit-action forms.

7.5.1. Must you smoke?

We start with type 1 implicit-action forms, those for which the action is determinable from a combination of the choice of indirect utterance form and the action named. REQUESTs do not seem to have any frozen forms of this type, but there is a pattern with a secondary REQUEST force, 7.12.

7.12 Must you <action>?

Paraphrases of 7.12, e.g., "Do you have to <action>?" are also type 1. The use of 7.12 indicates that a stopping or avoiding action is desired, and in particular the action to be stopped or avoided is the one named. For example, "Must you talk so loud?" is a request for either the implicit action "Stop talking so loud" or "Avoid talking so loud."

Example 7.12 and its paraphrases can be accounted for by the following rule:

Rule 7.1:[41]
P1 may ASK about P1's obligation to perform an action
-- instead of --
making a REQUEST by ASKing whether P2 can stop or avoid the action.
(Rule 6.5 acting on REQUEST precondition II.)

Rule 7.1, and the other rules that will be given for implicit actions, represents a link between an ISA pattern (the first half) and either a simple form or a pattern produced by one of the general rules from Section 6 (the second half).[42]

Example 7.12 and its paraphrases are the only type I implicit-action REQUESTs that I have found, but we have the SUGGEST forms "Why not <action>?" and "Why <action>?" For these forms, the suggestion is "do <action>" and "stop or avoid <action>", respectively.

The important feature of Rule 7.1 and the other rules in this class is that the choices of intended action (stopping or avoiding the action) are listed explicitly in the rule. Context plays a role in determining which of the two was intended by P1, but the operation for determining the implicit action is a choice, not an open-ended search.

7.5.2. Can you reach the salt?

For the next implicit-action form, type 2, P1 has some action in mind but the implicit action is not determinable from surface form alone. Type 2 implicit-action forms name

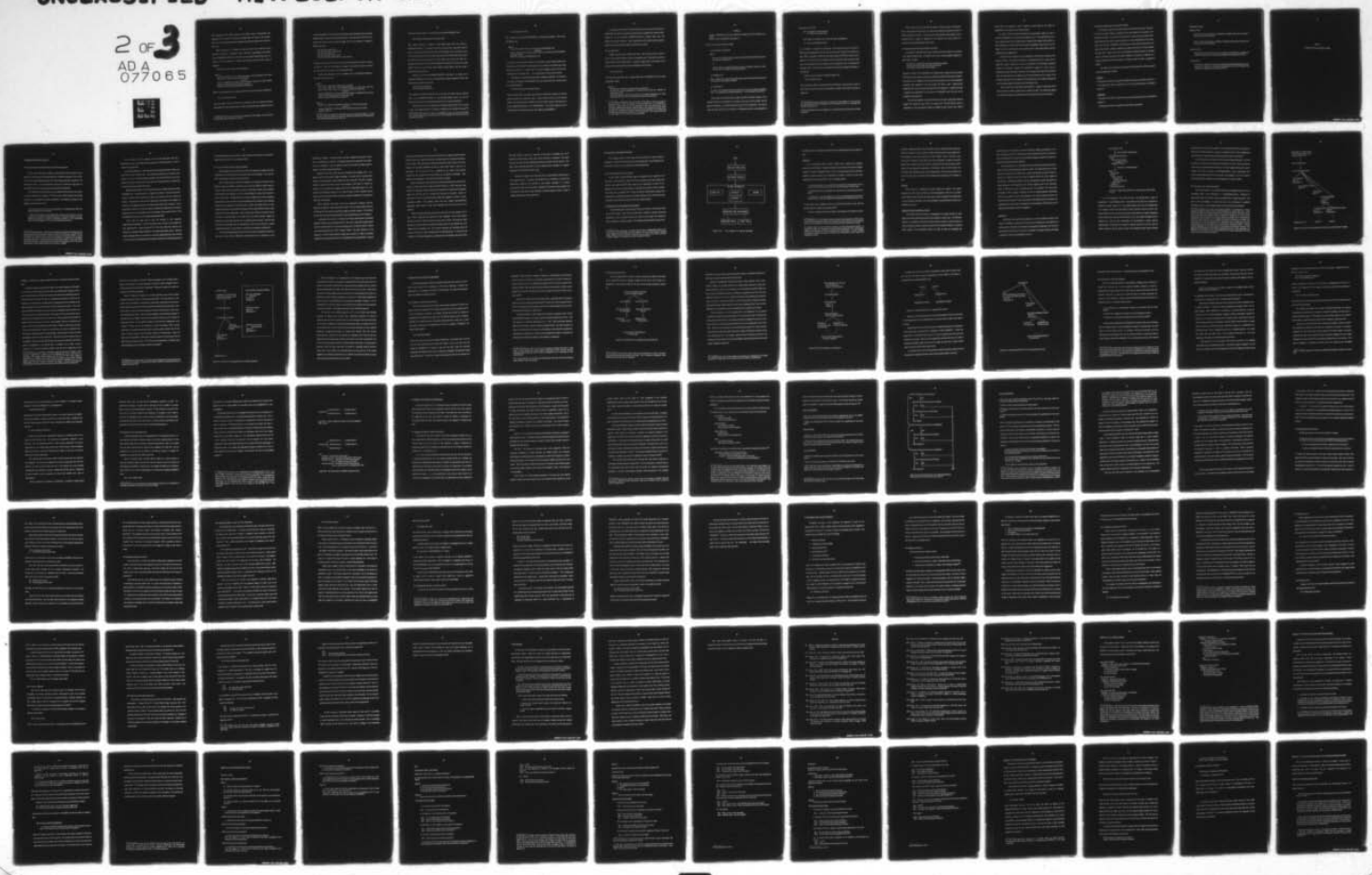
---

41. Earlier versions of Rule 7.1 and the other rules for implicit-action forms explicitly stated that the action intended by P1 should be clear to P2 from context. I now believe, however, that this condition is part of the more general constraint to avoid ambiguity, which applies to direct as well as indirect speech acts. (See Section 6.1.) It is therefore not necessary to restate the principle in each implicit-action rule.

42. Note that Rule 7.1 is explicitly tied to REQUESTs. It is not clear that implicit-action rules need to be tied so closely to individual speech acts. I have found some rules shared by pairs of action-centered speech acts, but further investigation in this area is needed.

AD-A077 065 MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTE--ETC F/G 5/7
TOWARD A COMPUTATIONAL THEORY OF INDIRECT SPEECH ACTS.(U)
OCT 79 G P BROWN N00014-75-C-0661
UNCLASSIFIED MIT/LCS/TR-223 NL

2 OF 3
AD A
077065



some component of the implicit action, and the implicit action is determinable from knowledge of the structure of actions plus knowledge about the context of the utterance. Context is necessary because the action component named may be associated with more than one action.

More specifically, for a definition of *component* I draw on the model of actions introduced in Section 4. A component is either a state that is a prerequisite of the implicit action, a PRINCIPAL-RESULT of the action, one of the action's semantic input cases, or a subaction.[43] In this subsection, we examine rules based on each of the four types of component.

We start with a type 2 rule based on prerequisites:

Rule 7.2:
P1 may ASK about or utter a representative about a stative prerequisite of an action where P2 is in some way a participant in the state
-- instead of --
making a REQUEST by ASKing about or uttering a representative about whether P2 can take responsibility for carrying out the action.
(Rules 6.5 and 6.6 acting on REQUEST precondition II.)

Rule 7.2 can be illustrated by an example of Searle's, the use of 7.13 to convey a REQUEST.

7.13 Can you reach the salt?

Here, the implicit action is that P2 pass the salt (either to P1 or to someone else clearly identifiable from the situational context). Instead of naming the action explicitly, P1 names

43. The subactions that can be used in type 2 implicit-action forms appear to be restricted to standard path steps (as defined in Section 4).

a stative prerequisite of the action. Note that naming a stative prerequisite does not always imply the action. One may ask "Is the salt near you?" or "Can you reach the salt?" for "Can you pass the salt?" but 7.14 or 7.15 do not imply 7.16 or 7.17 in response to a request to describe the dinner:

7.14 The salt was near Harry.
7.15 Harry was able to reach the salt.
7.16 Harry passed the salt.
7.17 Harry was able to pass the salt.
7.18 I asked for the salt and Harry was able to reach it.

Example 7.18 does imply that Harry passed the salt, but only because the speech act context was reported. Thus, the action implied is a property of the particular speech act context and not a property of the nature of stative prerequisites alone.

Turning from prerequisites, the next component type is the PRINCIPAL-RESULT. Two rules are of interest here:

Rule 7.3:
P1 may utter a representative about a desire or need[44]
for a state or object that is the PRINCIPAL-RESULT of some action, where the desire or need is P1's or occurs for someone P1 is empowered to speak for,
-- instead of --
making a REQUEST by uttering a representative about a comparable desire or need for P2 to take responsibility for carrying out the action.
(Rule 6.3 acting on REQUEST precondition 1.)

Rule 7.4:
P1 may ASK whether the PRINCIPAL-RESULT of a desired action obtains
-- instead of --
making a REQUEST by ASKing whether the desired action has been done.
(Rule 6.1, Clause iv)

---

44. *Desire* and *need* together are equivalent to *want* as I have been using it. I used "desire or need" here to make the rule read smoothly, but this phrase can be replaced by *want* without changing the sense of the rule.

An utterance related to Rule 7.3 is 7.19, which has an immediate REQUEST force.

7.19 I want this room clean when I come back, please.

This example illustrates a property of those implicit-action forms that involve a PRINCIPAL-RESULT. In uttering 7.19, P1 may or may not have a particular action in mind. On the one hand, P1 may want P2 to call the cleaning service, a particular action. On the other hand, P1 may wish to convey only a desire that P2 find some way to achieve the general goal "get the room clean." In this second case, P1 leaves the choice of a particular action to P2. Note that I will still consider this second use of 7.19 to have a single implicit action (e.g., "get the room clean"), keeping in mind, however, that the action is specified only in a general way.

Moving on to the next PRINCIPAL-RESULT rule, Rule 7.4 is similar to 7.3, producing REQUESTs such as 7.20. An explicit-action ISA that corresponds to 7.20 is 7.21, which is based on Rule 6.1, Clause (iv).

7.20 Do you have the letter?
7.21 Did you pick up the letter?

The component relationship should be clear in this pair of examples: having a physical object is the PRINCIPAL-RESULT of a pick-up-physical-object method. In turn, the picking up action done is the principal intended effect of an unuttered REQUEST.

We come now to the third componential relationship, semantic input cases of actions named. An example utterance is 7.22, spoken as a REQUEST that P1 light P2's cigarette.[45]

45. If, on the other hand, 7.22 is uttered as a REQUEST to give P1 a match (or matches) then I would consider the relationship involved to be prerequisite and the utterance to be derivable from Rule 7.2.

7.22 Do you have a match?

Here, matches are one possible INSTRUMENT in a light-cigarette method. Thus, we get the following rule:

Rule 7.5:
PI may ASK whether P2 has <item that fills an INSTRUMENT slot>
-- instead of --
making a REQUEST by ASKing whether P2 has <item that fills an INSTRUMENT slot> for/to <action>
(Rule 6.10 acting on the INSTRUMENT case)

It is not clear to me what other semantic cases may appear in type 2 implicit-action forms. Certainly, type 2 ISAs based on semantic input cases are similar to the explicit-action forms discussed in Section 6.3.2. It appears, however, that type 2 implicit-actions are derived from a more limited set of semantic cases. This is a question that I leave to further research.

This brings us to the fourth, and final, type of component, subactions. An example is 7.23, which in the proper context REQUESTs not only a transfer but also a creation action (i.e. write the program).

7.23 Give me a program that builds an arch.

Creation actions bring a new object, either physical or mental, into existence, and creation actions performed for someone else have a transfer step near the end of the standard path(s). In this transfer step, the AGENT of the action transfers control of the object that is in the PRINCIPAL-RESULT to the BENEFICIARY. This transfer subaction may be named instead of the creation action in a type 2 implicit-action REQUEST.

To summarize what has been said about type 2 forms, we can say that these forms are recognizable with reference to general knowledge about the components of different actions. *Components* are the prerequisites, PRINCIPAL-RESULTs, semantic input cases, and subactions described in Section 4. Type 2 implicit-action rules include some that apply to actions in general and some that apply to specific classes of actions.

7.5.3. It's cold in here.

Finally, we come to type 3 implicit-action forms. Type 3 forms are those in which a state or action named constitutes a basis for the implicit action. The notion of a basis for action will be discussed further below. A standard example of type 3 forms found in the literature is 5.19, repeated below, used as a REQUEST to close the door.

5.19 It's cold in here.

The rule that characterizes type 3 implicit-action forms for REQUESTs can be stated informally as follows:

Rule 7.6:
P1 may utter a representative about a basis for action where
P1 wants P2 to take responsibility for carrying out an action that remedies an undesirable state[46]
(This state constitutes the basis for action or has a special relationship to it. More specifically, the state is "S1" in the definition of a basis for action, below.)

---

46. In uttering a REQUEST derived from Rule 7.6, P1 reflects a desire to remedy an undesirable state. Thus, the negation of the undesirable state is a PRINCIPAL-RESULT desired by P1. In the last subsection, I said that implicit-action forms involving a PRINCIPAL-RESULT could be uttered with or without a particular action in mind. We see this quite clearly for utterances derived from Rule 7.6, as evidenced by the following common sort of ambiguity: P1: "It's cold in here." P2: "What do you want me to do about it?" P1: "I don't care. Just do something."

-- instead of --

making a REQUEST by uttering a representative about PI's desire or need for the action.
(Rule 6.3 acting on REQUEST precondition 1.)

A *basis for action* may be defined as follows:

i. a state (S1) that is undesirable

either:

because it is in conflict with some goal in the context (from the local context to the very general "social context")

or

because when S1 is combined with other states in the context it results in a state S2, where S2 is in conflict with some goal, as above

ii. a symptom of S1

Here, a *symptom* is a state S3 that typically coexists with S1, where when S3 occurs one is warranted to check whether S1 obtains.

iii. a signal for S1

A *signal* is the communication step of a process that is set up to monitor something else -- e.g., another process, a physical location, etc. -- for the existence of some state.

This is quite a bit of material; some examples may help to clarify the definition. First, example 5.19 above can be thought of as reporting a state that is in direct conflict with the goal of maintaining a comfortable temperature (part (i) of the definition). This type of basis for action can be used as an answer to a why-question asked by a non-participant in

the speech act, e.g. 7.24.[47]

7.24 P1: Why did you close the door?
P2: Because it was cold in here.

An example of a symptom used as a basis for action is the following:

7.25 The cake is beginning to smell.

A cake smelling is a symptom of its being done. The fact that the cake is in the oven (a contextual state) combined with the fact that the cake is done can result in a burned-cake state, which is generally in conflict with the goal of baking a cake.[48] Two remarks can be made here. First, while this sort of analysis of states can always be done, at least in many cases I do not believe that it must be done in order to recognize an ISA. I will expand on this point below. Second, note that a symptom need not name an unpleasant situation, as 7.25 illustrates.

Finally, we come to signals, an example of which is 7.26.

7.26 The buzzer went off.

The oven buzzer could be a signal for the fact that the cake is done, and, by a chain of states similar to that for 7.26, this state could lead to a conflict with the goal of having an edible cake.

---

47. I specify non-participant here because I suspect that, if the question in 7.24 is asked by the participant making the REQUEST, then the preferred answer is something like "Because you told me to."

48. This combination of states exemplifies the second half of part (i) of the basis for action definition.

There is still more to be said about the notion of a basis for action. In particular, there is the question of how a basis for action can be related to the general structure of actions developed in Section 4. Since the answer to this question involves quite a bit of additional detail, it appears in Appendix D. This permits us to move directly to a discussion of the special properties of type 3 forms.

7.5.4. Special Properties of Type 3 Implicit-Action Forms

In Section 7.3 I observed that the type 3 implicit-action class constitutes an exception to the single indirection limit. In addition to this, there is another property displayed by type 3 forms. Compare:

7.27 Get me a cup of coffee, and I'd like something to eat, please.
7.28 Get me a cup of coffee, and do you have any pie?
7.29 *Get me a cup of coffee, and I'm hungry.

Example 7.27 shows a frozen REQUEST form conjoined with a simple form, and example 7.28 contains a type-2 implicit-action form accounted for by Rule 7.2. (Not all type 2 implicit-action forms are as acceptable as 7.28, but all that I have looked at are at least marginally acceptable when conjoined in this way.) Finally, 7.29 shows a type 3 implicit-action REQUEST that cannot be successfully conjoined with the simple form. It appears that type 3 implicit-action REQUESTs are the only sort of REQUEST forms that may *not* be conjoined with simple forms.

The uniform possibility of double indirection and the special conjunction behavior support the treatment of type 3 forms as a separate class. We still, however, need an explanation of why such behavior should be exhibited by type 3 forms at all. While I

cannot answer this question in detail, I suggest an account based on the degree of independence of the speech act forces of an utterance.

For type 3 forms, the message conveyed by the representative speech act force is relatively independent of the action-centered speech act force. The immediate force of a type 2 implicit-action form, for example, names a *component* of the implicit action, which is a close relationship. In contrast, the "basis for action" relationship in type 3 forms is one in which the state or action named is more separate and on a more equal footing with the implicit action (see Appendix D).

To the extent that a speech act force is independent it can be seen as carrying its own separate message. The message conveyed by a representative force of a type 3 ISA is apparently separate enough that it merits the possibility of the additional modulation afforded by indirection. Similarly, the representative message of a type 3 form is independent enough of the other force(s) that, when conjoined with a type 3 action-centered form, the representative force competes with the action-centered force for a place in the semantic interpretation of the conjunction. (The difficulty here is that simple representatives do not normally conjoin with simple action-centered speech acts.)

Type 3 forms, then, exhibit special surface behavior. I suggest an explanation of this behavior based on semantic properties used to define ISA classes. The explanation appeals to a notion of the degree of independence between the speech act forces of an utterance.

7.6. Summary: Implications for the General ISA Rules

The categories given divide utterance forms into direct, indirect, and tag. Utterances are said to be single-, double-, or triple-force, and a given speech act may appear as the immediate, secondary, or tertiary force of the utterance. Frozen ISA forms are those forms whose immediate force differs from the speech act in a literal interpretation of the surface form of the utterance. REQUEST forms (and those of other action-centered speech acts) specify actions either explicitly or implicitly, and implicit-action forms are further classified into three types. Implicit-action forms are distinguished according to whether the intended action is derivable using surface form (type 1), whether it is named by a component of a method associated with the action (type 2), or whether a "basis for action" is appealed to (type 3).

The impact of this taxonomy on the general ISA rules (Rules 6.1 to 6.10 from Section 6) can be summarized as follows:

Additions

1. The general ISA rules are augmented by implicit-action rules as described in Section 7.5.

2. The general ISA rules are augmented by a set of frozen form patterns as described in Section 7.4.

Limitations

1. The general ISA rules are amended to permit only simple speech act forms as discussed in Section 7.3.

2. The general ISA rules are applied once only, except as noted below.

Production of Forces

Immediate Force

Frozen ISA Forms: produced by a combination of implicit-action rules as necessary and the patterns for frozen forms

Others: directly derivable from a combination of implicit-action rules as necessary and the surface form of the utterance

Secondary Force

All ISAs other than frozen forms have a secondary force derivable by a combination of implicit-action rules as necessary plus one application of the general ISA rules.

Tertiary Force

Of the forms examined here, only type 3 implicit-action forms (uniformly) may carry a tertiary force. This force is derivable by a combination of the type 3 implicit-action rules plus two applications of the general ISA rules.

PART 2

COMPUTATIONAL IMPLICATIONS

PRECEDING PAGE BLANK-NOT FILMED

8. Computational Perspectives on ISAs

8.1. Drawing Computational Implications from the ISA Categorization

The first part of this paper developed a characterization of indirect speech acts and justified it with traditional (non-computational) sorts of linguistic arguments. Now, in Part 2 some of the computational implications of this ISA categorization will be considered. The main point of Part 2 is that the distinctions made in Part I between ISA types occur at precisely the places where computational processes should be distinguished.

The computational implications that I wish to explore will be presented in the context of a process model for ISA recognition. [49] That is to say, I will be concerned with relating utterances to an appropriate underlying representation. The following two questions will, together, provide a focus for Part 2:

1. Given an utterance form, how can one represent, in a computationally useful way, the possible speech acts it can realize?

2. How can the system for characterizing ISAs from Part I be related to a view of recognition as essentially a path-building process? The paths in question are links between the speech acts conveyed by an utterance and nodes on a tree representing the events of the dialogue. (This view will be developed in Section 9.)

---

49. I use the word *recognition* rather than *understanding* to emphasize the importance that expectations will play in the process model; see Section II. Note that only recognition will be considered here. I see generation as at once an easier and a more difficult problem. On the one hand, the speaker is spared some of the problems of ambiguity and incomplete knowledge about what is said, making the generation process simpler. On the other hand, the generation process is complicated by a need to take the recognition process of the hearer into account, to avoid ambiguity when this is necessary or desirable.

PRECEDING PAGE BLANK-NOT FILMED

For the answer to the first question, I draw on the distinctions from Part I. Depending on category, the model will use explicit lists of realization patterns or rules at various levels of generality.

In answering Question 2, I will be concerned with the type and number of paths to be constructed for each utterance form. The immediate/secondary/tertiary distinction from Section 7.3 is seen as directly affecting this question, with one path built for each force. Another factor here is the direction in which paths are constructed; choice of direction has an impact on the representation of ISA forms.

Although these questions were stated separately, they are actually strongly interrelated. Accordingly, after a general discussion of recognition in Section 9, I discuss the questions together in Sections 10 and 11 and go on to look at some aspects of the role of context in Section 12. Note that the answers to the two questions are not intended to be complete. In no sense am I attempting to present a full computational model of ISA recognition. I have concentrated on what I see as the more important computational implications of the theory presented in Part I, with the hope that the ISA recognition scheme outlined in Part 2 will eventually form the basis of a more complete computational model.

Every effort has been made to make the discussion of ISA recognition implementation-independent. It is worth mentioning, however, that parts of the model have been implemented in a system described in [3]. This system, called Susie Software, was designed as a prototype dialogue module for an automatic programming system. While the system's debugged behavior was limited to a twenty line dialogue (plus some close variants), it is fair to say that the implementation did nothing to disconfirm the model presented here.

The implementation process did, however, reflect negatively on some earlier versions of the model, and in that sense it was an important exercise.

8.2. A Look at Related Work in Dialogue Modelling

Before plunging in, it may be useful to see where the approach to dialogue (and ISA) recognition that I am advocating fits in with other recent work on dialogue. I start with the OWL-1 representation of actions discussed in Section 4.

OWL-1 methods represent chunks of knowledge in the tradition of Minsky's frames [19] and Schank and Abelson's scripts [27], and as such they should be familiar types of structures. Both of these approaches are geared to representing common sense knowledge, frames primarily from the point of view of visual knowledge and scripts from the point of view of social actions. Although Minsky does discuss the application of frames to language understanding, he does not address the special problems posed by dialogue, where language generation and recognition are intimately related. Schank and Abelson's focus on social actions brings them closer to problems of dialogue, but their representational scheme is quite different from the one adopted here. Scripts are built from a small set of primitive actions, in contrast to the open-ended set of actions used in the OWL-1 methods. Despite the differences, the three approaches have one very important point in common. All three focus on chunks of knowledge whose interrelationships are used to form *expectations* to guide recognition processing. This idea plays a central role in the framework described here.

Of the specifically language-related research, four efforts are especially relevant here. The first is work by Grosz [8][12], who focussed attention on the importance of task

knowledge for dialogue. A semantic network was used to represent the structure of tasks such as assembing an air compressor, and dialogue utterances were interpreted with respect to this task knowledge. Using this approach, Grosz has been able to develop important insights on reference and pronominalization.

Although Grosz's work shows how task knowledge affects dialogue, there is no explicit model of more general dialogue knowledge. One early effort to represent both linguistic and non-linguistic actions uniformly was that of Bruce [5]. The OWL-1 dialogue method is quite close to Bruce's social action paradigm, in that both are centered on semantic cases and both were developed with an eye to incorporating speech acts. The primary representational difference is the distinction made in OWL-1 methods between standard and recovery paths (Section 4), which has important implications for processing as well as for representation.

Bruce's approach, then, points the way to an integration of linguistic and non-linguistic knowledge in a single representation scheme. There is still the problem, however, of showing how such general representations apply to particular situations. (In [5], Bruce discusses the use of social action paradigms for story understanding but not for dialogue.)

A group that has worked on a mechanism for relating representations to individual dialogues is Moore, Levin, and Mann (see [20]). Their work is specifically dialogue-related, and their representations for what they call *dialogue games* are spiritual cousins to Bruce's social action paradigms and to OWL-1 dialogue methods. The major difference is in the approach to control structure; general dialogue game structures are related to particular dialogues by a group of processors that work independently and in parallel [16]. In contrast,

the design that will be discussed in Section 9 and Section 11 has a highly centralized control structure where every effort has been made to identify limited sets of information necessary for each choice. This means that an attempt has been made to limit searches and match attempts by structuring information so that the system knows where to look for its information. To an extent, this view is influenced by the current serial machine architectures, but I suspect that the more structure we can identify in knowledge -- both linguistic and non-linguistic -- the more successful will be our models.

Another group that has been concerned with both representation and mechanism for dialogue is Cohen, Perrault, and Allen (Cohen [6], Perrault et al. [22]). They have been interested in specifying a general, inference-based mechanism for choosing and recognizing speech acts, given the facts of a situation and the goals of the speaker. Speech act generation is viewed as a general planning problem, and recogition is viewed as a plan-recognition problem. This approach differs from more strongly expectation-based approaches, which include the other efforts described here as well as the one to be presented.

Within this plan-recognition framework, the work that is of most relevance for us here is that of Allen [1], who concentrates on ISA recognition. Of the three theories discussed in Section 5, Allen's framework is closest to that of Searle. The model that Allen presents uses a general inference mechanism to identify speech act forces beyond what I am calling here the immediate force. The inference mechanism uses knowledge about the structure of actions, as well as knowledge about the planning process. As noted above in Section 6.5, such a general approach is probably necessary for handling some types of ISAs.

My claim, however, is that, for a large class of ISAs, there is knowledge that can be exploited to avoid deriving a form "from scratch" each time it is recognized. The design discussed here relies heavily on predigested information associated with ISA patterns. This allows more aggregated operations, so that one operation may correspond to a segment containing several inferences in Allen's system.

To sum up, the approach that will be taken gives a central position to expectations in the recognition process. It recognizes the importance of task knowledge in framing these expectations, integrating this knowledge in with linguistic knowledge in a uniform representation. Efforts have been made to consolidate control structure where possible, and an attempt has been made to tailor operations to different categories of ISAs, in order to exploit knowledge that we have about them.

9. A Framework for Recognition Processing

The recognition process for ISAs cannot be discussed outside of a general model of recognition. In Section 9.1 I briefly discuss the system configuration I am presupposing, and in Section 9.2 I go on to recast recognition as a path-building process.

9.1. A System Configuration for Recognition

In this section, I discuss only those aspects of recognition that are required for the discussion of ISA recognition. Figure 9.1 shows the recognition processing configuration that the ISA model assumes.[50] Many aspects of the configuration in Figure 9.1 are influenced by the wish to model dialogue as, essentially, the execution of OWL-1 methods. This means that a knowledge base is needed to hold the method library (among other information), a knowledge base processor is needed to maintain it, and a method interpreter is needed to move from step to step and invoke and terminate methods.

Knowledge Base and Knowledge Base Maintenance

The Knowledge Base contains concepts, the basic unit of OWL-1, and structures built out of these concepts. These structures include OWL-1 methods to represent processing knowledge as well as descriptive structures for classes of static concepts (e.g. substantives such as toy blocks or computer consoles). The Knowledge Base Maintenance facility needs

---

50. Figure 9.1 has many points in common with the OWL-1 implementation done at the M.I.T. Laboratory for Computer Science by William A. Martin, Lowell Hawkinson, William Long, William Swartout, Alexander Sunguroff, and the author. That system contained running examples of everything in 9.1 except the English Front End.

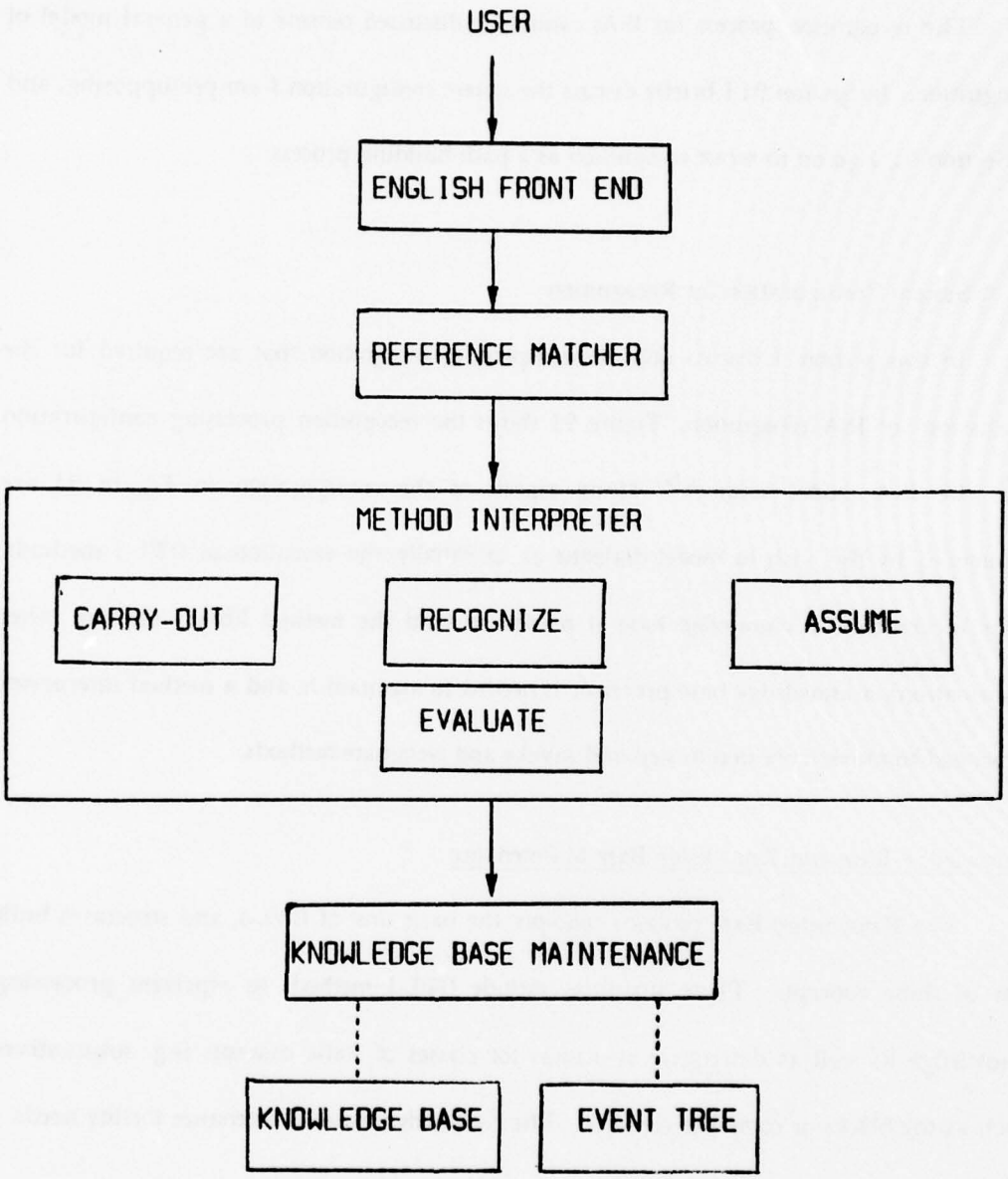


Figure 9.1. The recognition system assumed

the ability to add new concepts and structures to the Knowledge Base and to modify old ones.

Interpreter

It is the Interpreter's job to "execute" method steps to generate and recognize utterances in the dialogue. If request-and-respond, ask-and-answer and the other dialogue methods are speaker independent, then it is up to the interpretation process to determine whether a particular utterance is to be generated or recognized. There are, in fact, three possible modes of interpretation[51] for a step in a dialogue method:

1. Carry out the step (e.g., ask a question). In dialogue, the most important role of Carry-out is to generate utterances, but it is also used to execute "mental" actions.

2. Recognize that a step has happened (e.g., that an answer to your question has been given).

3. Assume that a step has happened (e.g., if your conversational partner gave the answer, then he or she had to perform the mental process of finding the answer first.)

A set of simple rules is adequate in most cases to determine the mode of a step from the semantic input case settings in the procedure call.[52]

In order to introduce other system modules, I will expand on the Interpreter's Carry-

---

51. Throughout Part 2 of this paper, the words *interpret* and *interpretation* will be used to indicate the interpretation of methods. The meaning of *interpret* found in *natural language interpretation* (as opposed to *natural language generation*) will be conveyed by *recognize*.

52. Semantic input cases were discussed in Section 4. In some conversational settings it is ambiguous whether, say, recognition or merely assumption should occur next. "Constructed" dialogue environments can avoid this problem by careful design. In other dialogue environments, clarifying discussion can often be used to determine the currect mode in an ambiguous situation.

out mode. (Recognition mode is discussed further in the next subsection.) Stative steps (e.g. results) are handled in Carry-out mode by simply asserting them into the Knowledge Base. The handling of non-stative method steps is more complex. Given a non-stative step, Carry-out first evaluates it with respect to the current environment. (Evaluate is discussed below.) This evaluated step is then treated as a procedure call and matched against known methods to find appropriate subprocedures. Restrictions on what can fill each input case associated with a method are used in this matching process. The use of the evaluated call means that the choice of a method is heavily influenced by the progress of the dialogue to date.

Evaluate

OWL-1 forms are a combination of regular concepts and variables. The module Evaluate takes these forms and looks up and binds appropriate values of the variables. The evaluation process is always done with respect to some environment, for example when a call is evaluated with respect to the current dialogue environment before the search for a method to carry it out.

English Front End and Reference Matcher

The English Front End consists of morphological and parsing modules for typed input and includes additional signal processing modules for spoken input. The Reference Matcher is not limited to resolving noun group references; instead, the entire utterance is seen as referring to semantic structures, and the task of the Reference Matcher is to identify these concepts. This identification process can either be done by composing new

representations, by verifying that the utterance matches an expected representation, or by some combination of the two. Note that the arrows in Figures 9.1 and 9.2 are intended to mark the dominant flow of control. Some interaction with adjacent modules is expected for both the Front End and Reference Matcher.

For our purposes here, it is the output of the Front End and the Reference Matcher, rather than their internal operations, which is of interest. The representation that we need from the Front End will be called surface semantic representation, so-named because its elements are semantic structures but its form corresponds closely to the surface English of the utterance. The Reference Matcher takes this representation as input and produces as output Interpreter level representation, the semantic language used by the Interpreter. Figure 9.2 illustrates the sorts of representations that I am assuming. The lexical items correspond to OWL-1 concepts. Parentheses are used to establish grouping relationships; they have no relation at all to LISP parentheses. The primary differences between the representations in Figure 9.2 is that, where possible, referent identification and semantic case assignments have been made in the Interpreter level representation. Further details of the representations assumed are given in Appendix E.

Event Tree

The Event Tree is built by the Interpreter as a record of methods executed in the course of a dialogue. In this regard, the Event Tree can be thought of as intermediate term memory, used to record the current dialogue and organized chronologically. Past events are not removed from the tree, so that they are available for inspection, question answering, resumption (in the case of uncompleted events), etc.

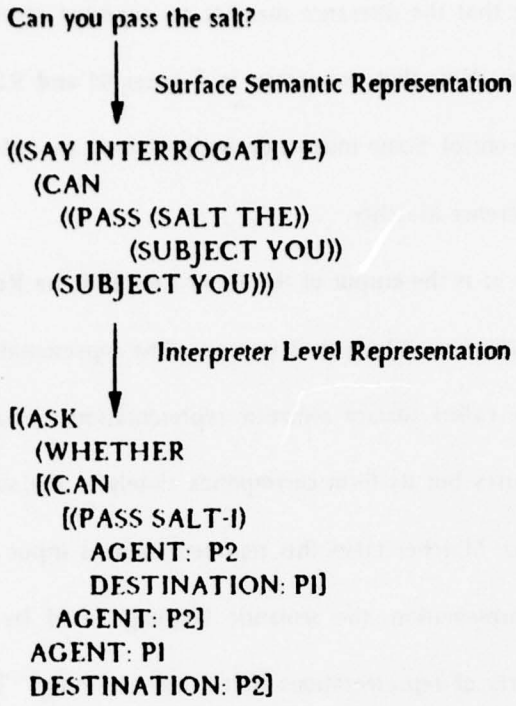


Figure 9.2 Example representations for an utterance that conveys (only) an ASK.

It is to the branches of the Event Tree that I am referring when I speak of recognition as a "path-building" process. The branches of the Event Tree correspond to substeps of methods. Figure 9.3 is a very simple example of an Event Tree configuration for a computer console session environment in which a question is asked by the user and answered by the system. Note that each node on a branch either has a subpart relationship with its superior (e.g., ASK and ask-and-answer) or it is a more specific description than its superior (e.g., "find out some information" and ask-and-answer). Surface semantic representations of utterances do not appear as separate nodes on the tree, but they are instead associated with the speech act node. The association between surface semantic

representation and event node corresponds to the relationship between ISA and speech act. Note that the example in Figure 9.3 is quite a bit simpler than most. Once a dialogue gets underway there are typically many levels of subcalls on the Event Tree branches.[53]

Two of the speech acts in Figure 9.3 have not been introduced previously. The first, STATE, is a representative that conveys what PI believes to be a fact. Quite a bit must be said in order to adequately define the notion of a "fact" but I appeal instead to the reader's intuition. The other new speech act, ACKNOWLEDGE, is also a representative that conveys information about PI's status with respect to the dialogue. It is a sort of marker, indicating that PI has completed a step in the dialogue and is ready to continue.

9.2. The Event Tree and Recognition Mode

This section contains a very condensed outline of the recognition process that I am advocating, which I have characterized as a "path-building process." Although the information content of this section is dense, everything relates to a single major point. This point is that path building is a process that occurs in an environment of incomplete knowledge, so that it is computationally important in what direction path building is

---

53. There is another way in which the treatment described is a simplification. In many dialogue situations, utterances occur whose motivations are learned only later. These motivations could have an effect on the choice of some of the higher level nodes on the Event Tree path of the original utterance. (The level of concern here is the "get some information" node in Figure 9.3.) The usual options are available to the system builder: attaching a path to the tree can be deferred until further information is available, the system can guess and back up, or minimal commitments can be made (e.g., "get some information" can be chosen temporarily as a motivation). If the last alternative is chosen, the descriptions on the event path can be refined as more information comes in. For the dialogue modelling task, this third alternative is probably the most suitable in general. This approach avoids problems of non-uniform representation that occur when attachments are delayed, and it avoids the need to undo premature and erroneous commitments that may occur with a guess and back up strategy.

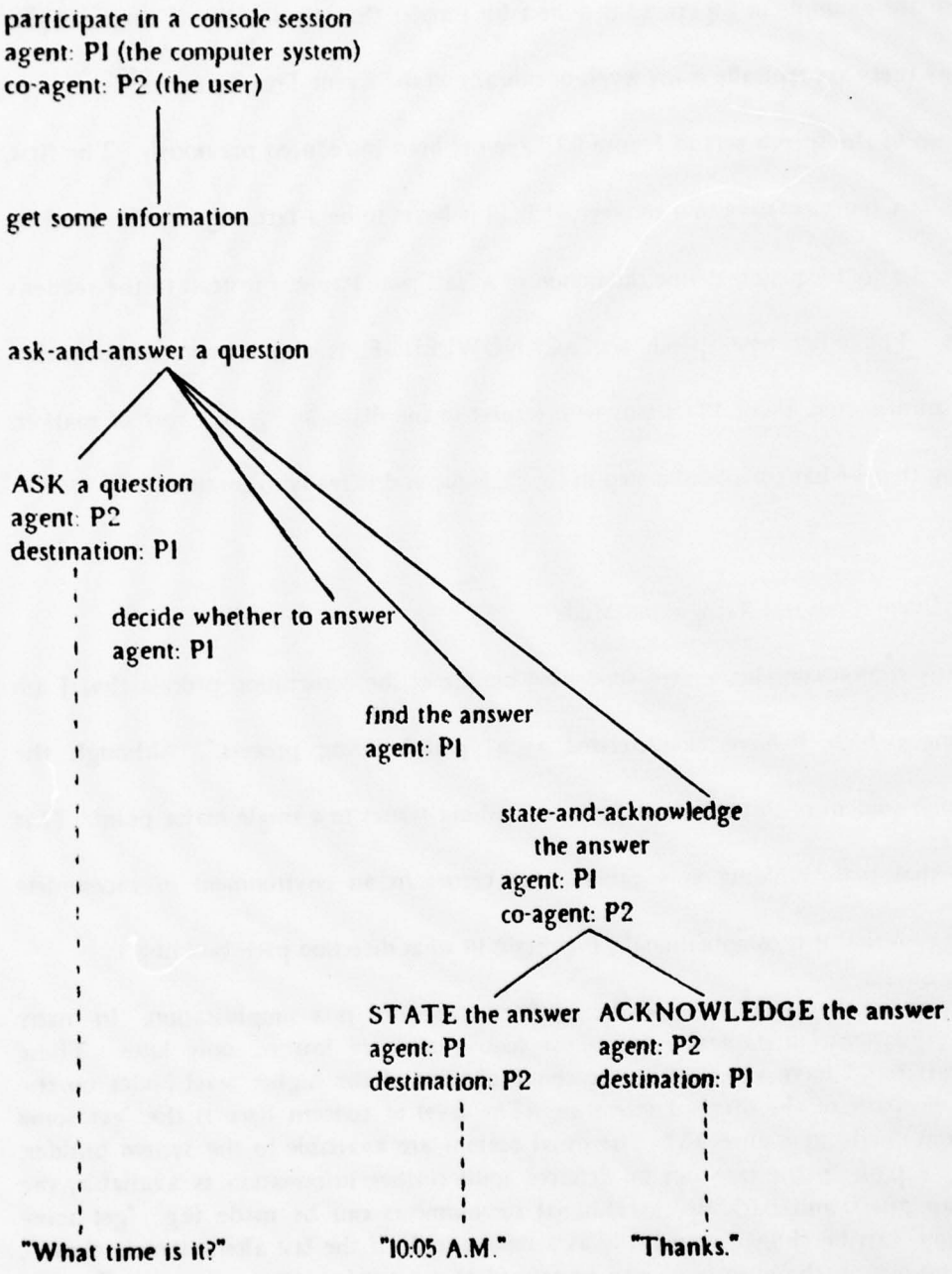


Figure 9.3 An Event Tree configuration with paths for a question/answer exchange.

attempted. I outline here a suggested approach based on a distinction between utterance types.

Recall that methods are being used to model a task-oriented dialogue from the point of view of one of the participants (say, P1) so that some method steps correspond to all or part of the underlying form of generated utterances and some provide underlying forms for utterances that are recognized. Once P1 uses a method step to generate an utterance we can identify a set of patterns that are highly likely as underlying forms of the response by the other participant P2, given that the dialogue continues along the same lines. This set of patterns includes the possible active successor steps to the generated step, the lead-ins (i.e. first step executed in recognition mode) to recovery paths related to the generated step, or lead-ins to recovery paths related to its successor steps. (Recovery paths were introduced in Section 4.) Here, *active* is used to rule out strictly conceptual steps such as thinking of an answer. Since P2 need not respond only to P1's goals, I include two other types of pattern, initiator and metadiscussion. Initiators are utterances that start off new, independent tasks (as opposed to substeps of tasks currently underway). Metadiscussion utterances are about the dialogue rather than strictly step-derived. They tend to clarify or change the flow of the dialogue. (For more on these different types of utterances, see Brown [4].) The different patterns corresponding to the different types of utterances will be called *structural expectations*. [54] Structural expectations developed from the current configuration of the

54. The term *structure* is used here for what others might call the syntax of dialogue. Many "structural" phenomena are semantic in flavor (although they do not necessarily vary according to the specific semantic domain), and the use of the term *syntax* might be misleading. For example, the fact that questions get answers can be called "dialogue syntactic;" however, the fact that answers may have associated stipulations ("Yes, if...") or qualifications ("Yes, but...") begins to stretch the term *syntactic*. The expectations generated from methods and the general task environment are therefore referred to as structural.

Event Tree can be viewed as "on deck", waiting to participate in the matching process. When a match against a structural expectation is successful, a path of dialogue events is created, linking nodes for utterances to expectations. This path is attached as a branch on the Event Tree (of this, more below).

Figure 9.4 illustrates the notion of a structural expectation, again with the simple example of a question asked in a console session environment. This time, however, it is the system that asks the question. If all proceeds according to the standard path of the ask-and-answer method, the user will first find the answer (an assumed step from the point of view of the system). As the ask-and-answer method is written now, the next step is a call to state-and-acknowledge the answer. Thus, state-and-acknowledge is one expectation, although the Interpreter must go through a layer of calls and method searches to find the speech act expectation STATE. It is the STATE step that is at the same level of aggregation as the utterance. [55] Since the user may decide not to answer the question, other structural expectations are possible, among them lead-ins to recovery paths associated with ask-and-answer, lead-ins to general recovery methods, initiators, and metadiscussion. Figure 9.4 shows one such possibility, a STATE that is the lead-in step to the recovery path that says that P2 does not know the answer. (This STATE is not represented as a substep of state-and-acknowledge, because the expected responses are different.)

---

55. Additional mechanism, however, is needed to bridge the gap between the speech act (e.g., STATE) and the utterance. The sections that follow consider the mechanism necessary to bridge this gap for ISAs.

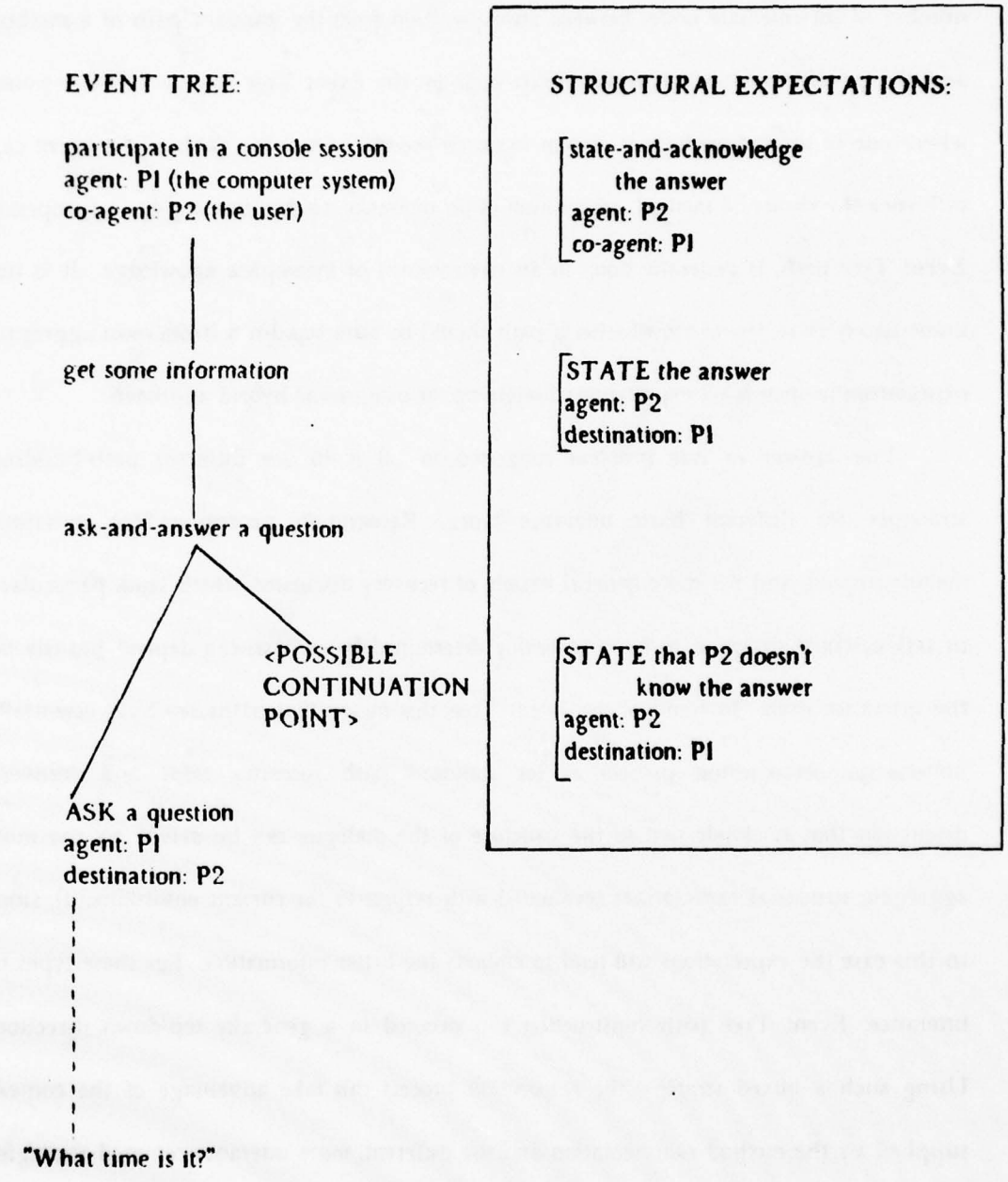


Figure 9.4. An Event Tree configuration with a structural expectation

Note that although it was straightforward in the example to get from "state-and-acknowledge the answer" to the speech act expectation STATE, in general there may be a number of intermediate nodes between an expectation from the standard path of a method and the possible speech act step(s). Each node on the Event Tree can be a choice point, where one of several methods is chosen to match the call. Since the goals of the agent can influence the choice of method, recognition of an utterance, i.e. constructing the appropriate Event Tree path, is generally done in an environment of incomplete knowledge. It is not immediately clear, therefore, whether a path should be built top-down (from most aggregate expectation to speech act expectation), bottom-up, or using some hybrid approach.

The answer to this problem suggested in [3] is to use different path-building strategies for different basic utterance types. Recognition procedures for initiators, metadiscussion, and the more general variety of recovery discussion which tend, particulary in task-oriented dialogue, to be only weakly determined by context, can depend heavily on the utterance itself. In terms of the Event Tree, this means that paths are built essentially bottom-up. Recognition procedures for standard path successor steps and recovery discussion that is closely tied to the structure of the dialogue can be driven by the most aggregate structural expectations (evaluated with respect to the current environment), since in this case the expectations will tend to embody the better information. For these types of utterance, Event Tree path construction can proceed in a generally top-down direction. Using such a mixed strategy, the recognition process can take advantage of the context supplied by the method representation and use different, more utterance-centered strategies when strong contextual information is not available.

10. General Observations about ISA Recognition

Given the general process model for recognition outlined in the last section, I can now discuss issues related to ISA recognition. In the first two subsections I consider the implications of the recognition framework for ISA processing. The final two subsections discuss the problem of ambiguity for ISAs.

10.1. The Role of the Recognition Framework

The structural expectations and the basic utterance types introduced in Section 9 can play a role in the recognition of ISAs. Within this framework, most of the examples used so far can be thought of as initiators. For example, "Can you pass the salt?" can initiate an independent task associated with eating a meal. Not all ISAs, however, occur as initiators. Consider, for example, the process of paying for merchandise with a credit card. If, some way along in the process, the salesperson utters 10.1, this is typically a REQUEST, not merely a representative.

10.1 I need your signature.

I claim that a good way for a process model to determine the correct speech act is to use the fact that the pay-with-credit-card method will contain a step to the effect that the salesperson REQUESTs the customer to sign the credit slip. Using this structural expectation, the Interpreter knows that it is looking for a REQUEST force (possibly among other alternatives). It can then ask whether the incoming utterance can be construed as a

REQUEST. More precisely, to recognize an utterance as a standard path or recovery path successor step or as a lead-in to a recovery path, the Interpreter can do what is necessary to find the speech act(s) (i.e. by hypothesizing possible Event Tree paths top-down from evaluated method steps), find the ISA forms associated with this speech act (either stored explicitly or in the form of general rules), and ask whether the utterance matches any of these forms.[56]

In contrast to standard and recovery paths, context is generally weaker for initiators, general recovery discussion, and metadiscussion.[57] In these cases, I have proposed that the Event Tree path(s) be grown bottom-up. For ISAs, this means starting with an ISA form and asking what speech acts it could convey.

In thinking about ISA representations and the details of recognition, then, I will be considering both directions: ISA form to speech act (henceforth abbreviated as ISA -> SA) and speech act to ISA form (abbreviated as SA -> ISA). This bi-directional approach follows from the choices made in the general recognition model. Since both directions are covered, it also means that the discussion in the following sections should be translatable to dialogue process models that are similar to the one proposed but which use a different mix of bottom-up and top-down strategies to relate utterances to expectations.

---

56. Here and elsewhere, when I talk in terms of utterances matching ISA forms, I really mean a match between the surface semantic representation of the utterance and the transitional representation of the ISA form. See Appendix E for a discussion of these representations.

57. It is not, however, clear to what extent metadiscussion uses ISAs, since these utterances tend to fall into rather fixed patterns.

10.2. ISAs and the Event Tree

In order to represent ISAs, one Event Tree path is built for each speech act conveyed by an utterance.[58] Given this approach, example 5.19 ("It's cold in here" uttered as a REQUEST to close the door) would have the Event Tree informally represented in Figure 10.1.

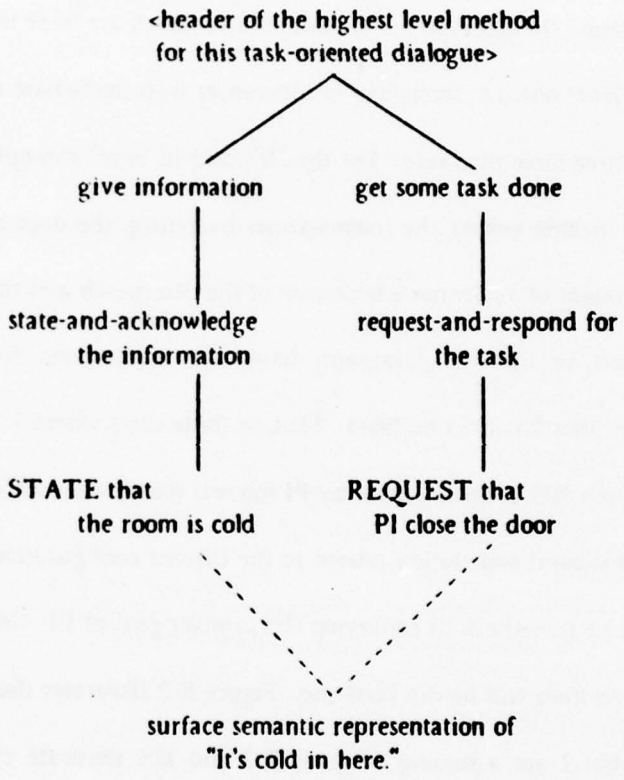


Figure 10.1. A possible Event Tree configuration for Example 5.19.

58. An utterance may, of course, suggest more than one possible set of Event Tree paths, and the choice between them may not be obvious. This is a problem of ambiguity, which will be discussed in the next two subsections.

Recall that the surface semantic representation of the utterance is associated with speech act nodes but is not itself an actual node on the Event Tree.

Despite the multiple path representation given to some utterances, ISAs in task-oriented dialogue are typically motivated by a single primary hierarchy of goals and subgoals held by PI. That is, in the usual case, *one* of the speech acts conveyed reflects the primary motivation behind the utterance. [59] Moreover, the speech act force that is the most important will be the *final* one, i.e., secondary for utterances with immediate and secondary forces, tertiary for a three-force utterance. For the "It's cold in here" example the primary goal hierarchy would include getting the room warmer by getting the door shut, and thus REQUEST can be thought of as the more important of the two speech acts conveyed.

The observations in the last paragraph have two implications for the general recognition mechanism that has been outlined. First, in those cases where I advocated top-down processing (Section 9.2), *if an utterance by PI conveys multiple forces, only one* of the forces will match a structural expectation related to the current configuration of the Event Tree. This force can be thought of as conveying the primary goal of PI. Second, the force that matches the expectation will be the final one. Figure 10.2 illustrates the type of Event Tree configuration that I am assuming. Figures 10.3 and 10.4 illustrate cases that I am assuming to be non-standard in task-oriented dialogue, because they contradict the first and second assumptions, respectively.

---

59. I emphasize that this restriction applies to the usual case. Exceptions are, of course, possible, but I would expect to see additional mechanism called in to handle them.

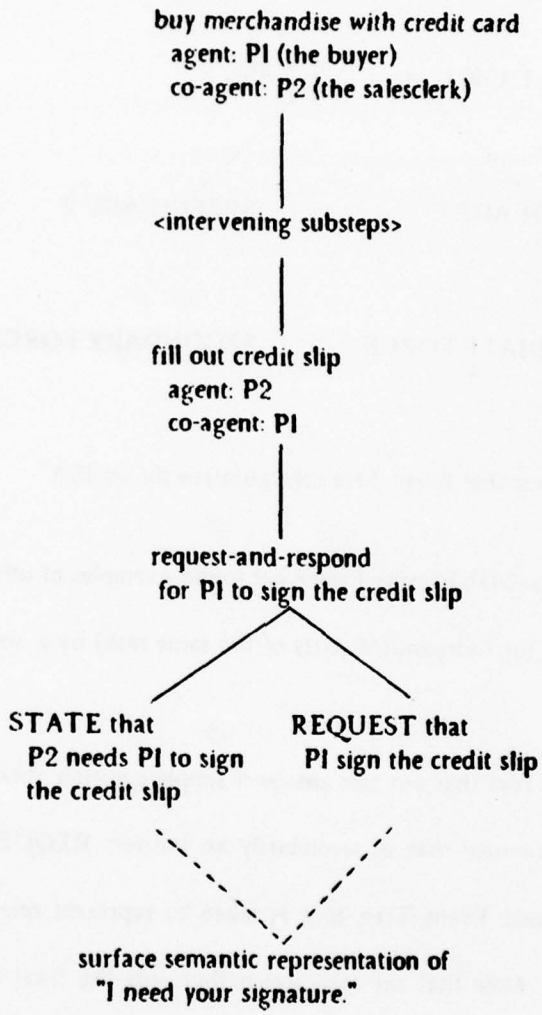


Figure 10.2. Event Tree paths for a two-force ISA.

In Figure 10.2, note that the STATE and REQUEST nodes would be marked with the same time and surface semantic representation, so there would be no chance of construing them as totally independent events.

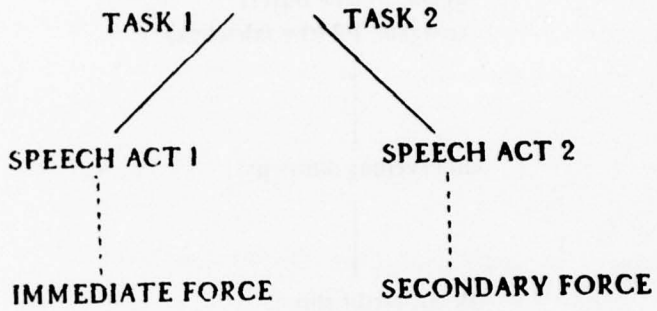


Figure 10.3. A questionable Event Tree configuration for an ISA

Figure 10.3 is only skeletal, because I have not found examples of utterances that relate to two independent tasks (or independent parts of the same task) by a split in their speech act forces.

Finally, Figure 10.4 says that one can answer a simple question about the temperature of the room with an utterance that is, secondarily an indirect REQUEST. I claim that Figure 10.4 is an ill-formed Event Tree, if it is taken to represent one of the ISA types discussed in this paper. Note that the assumption *that only the final force* can match a structural expectation related to the Event Tree implies that, once such a match is found, there is no need to derive further indirect forces.

In summary, Event Tree ISA representations involve one path per speech act force. For top-down recognition processing, however, only one path matches a structural

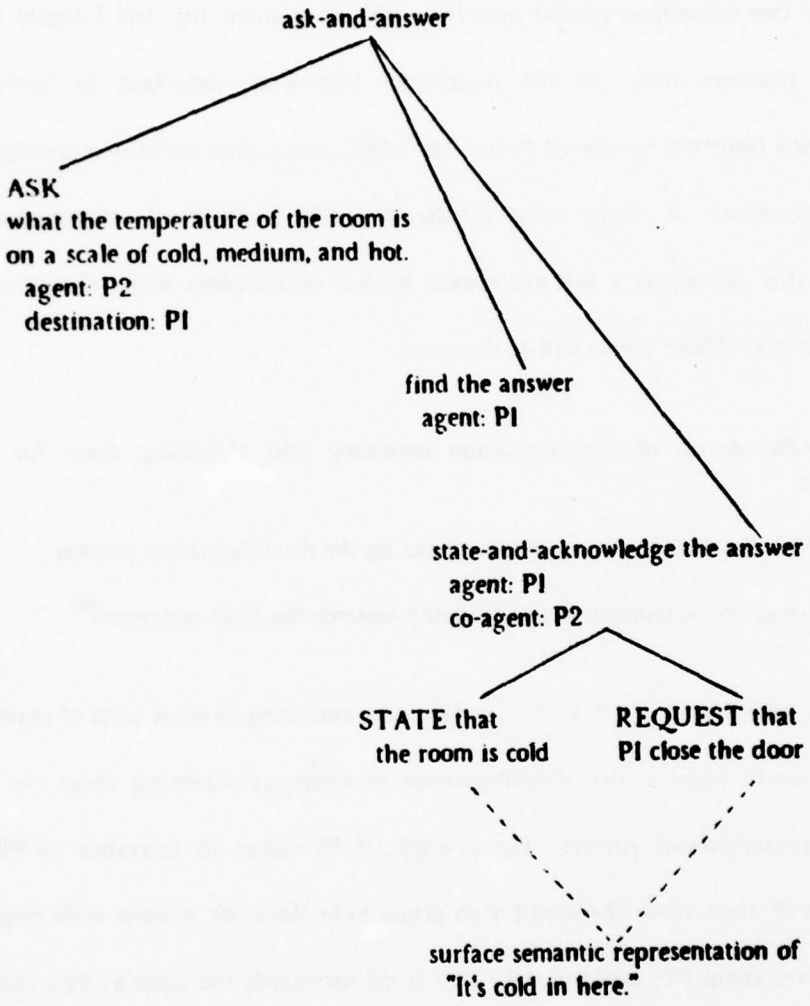


Figure 10.4. Another questionable Event Tree configuration for an ISA.

expectation related to the Event Tree -- the path generated for the final speech act force.

10.3. Some Remarks about ISA Ambiguity

The last two subsections glossed over the problems of ambiguity, and I would like to remedy that situation now. In the recognition framework described in Section 9, ambiguities with respect to speech act force occur when an utterance matches more than one structural expectation. A choice must usually be made as to which expectation was intended. In this subsection, I will make some general observations about the process of making this choice. Three points will be discussed:

1. the relationship of disambiguation reasoning and reasoning done for other purposes

2. the effect that a dialogue environment has on the disambiguation process

3. systematic versus nonsystematic ambiguity between the ISAs conveyed[60]

Starting with the relationship of disambiguation reasoning to other sorts of reasoning, the most important point is that disambiguation reasoning is reasoning *about the world model of the conversational partner*. For example, if P1 makes an utterance to P2, any reasoning that P2 does about the need for an action to be done, etc. is done with respect to what P2 believes about P1's world model. This is *not* necessarily the same as P2's model of

60. Searle in [30] has questioned the use of the term *ambiguity* to refer to differences in the forces conveyed by indirect forms. Since the notion of meaning that I am using (see Section 7.3) is broader than Searle's, and, since differences in conveyed forces amount to differences in meaning as I construe it, I will continue to use *disambiguation* to mean choices among the possible ISA forces of an utterance.

the world (or for that matter as P1's; this is a common source of error). This said, I will not say much more about the divergence of world models. Recent work in this area is that of Cohen [6] To justify limiting the scope of this paper, I appeal to what appears to be the default assumption:

> If there is no information to the contrary, assume that the dialogue partner's world model is the same as your own.

So, although I will note places where dual world models become a factor, I will assume for the purposes of exposition that there is "no information to the contrary."

Reasoning for disambiguation, then, does not always use the same knowledge sources as general reasoning. A further point about disambiguation reasoning is that it can be quite complex. Where REQUEST is a possibility as a conveyed speech act, it may be necessary for P2 to evaluate capabilities, willingness, authority relationships, schedules of activities, etc. This sort of evaluation can be an expensive process. Note, however, that the same information needed to disambiguate is often also needed to decide whether to comply with a directive or to determine whether information conveyed in a representative was previously known. Thus, a mechanism that can save its reasoning processes -- or at least save those steps and results known to be of potential interest later on-- can avoid duplicating what appear to be inherently expensive processes.

This brings me to my second point, which relates to the effect of the dialogue environment on the disambiguation process. One important aspect of dialogue is the fact that P1 is available to aid P2 in disambigating P1's utterances. Dialogue participants quite

commonly ask for clarification of intent, e.g., I0.2 to a "Do you know...?" question or I0.3 to a "Can you <action>?" form:

I0.2 Are you asking me or telling me?
I0.3 Yes. Do you want me to?

Another frequently used strategy is for P2 to make a disambiguating choice clear in a response. An example here is I0.4 in response to a "Can you <action>?" form interpreted as a REQUEST:

I0.4 Sure. I'll get started right away.

One result of this sort of response is that P1 can go on to alert P2 if P2's interpretation was not the one that P1 intended.

The dialogue environment, then, provides additional opportunities for aquiring disambiguation information and, if responses are constructed carefully, it provides checks to prevent misunderstandings. [61] Disambiguation in task-oriented dialogue may be difficult, but the options available effectively limit how hard a dialogue participant must work.

We turn now to the third topic of this subsection, the distinction between systematic and nonsystematic ambiguity in the speech acts conveyed by an utterance. A systematic ambiguity for ISAs occurs when all forms within an ISA subset display a particular ambiguity with respect to the speech act conveyed. Systematic ambiguities for ISAs include ambiguities that occur between direct and indirect immediate forces of an utterance. This type of ambiguity is common to all ISAs that do not contain special identifying

61. For a systematic approach to differences among different linguistic modalities, see Rubin [23].

characteristics (e.g., some frozen ISA forms as described in Section 7.4). A familiar example is whether 5.3 has ASK or REQUEST as its immediate force.

5.3 Can you close the door?

Another type of systematic ISA ambiguity is whether or not a given secondary force applies. This type of ambiguity is common to all ISA forms which may convey a secondary force, and the familiar example is whether 5.4 conveys a REQUEST in addition to the immediate ASK force.

5.4 Are you able to close the door?

Given what has been said, nonsystematic ambiguities are ambiguities that are not shared by other similar forms. The occurrence of nonsystematic ambiguity is more *fortuitous*. *Such an ambiguity typically arises from a special configuration of structural* expectations, e.g., for standard path successor steps. Either because of lexical ambiguity within the utterance or because the expectations are similar to each other, an utterance matches more than one expectation.

In discussing aspects of the process model, I will be concerned only with systematic ISA ambiguity. For this type of ambiguity, the structural expectations that match the utterance are all of the same basic type; that is, the utterance has several possible *interpretations* as an *initiator*, or several as a standard path successor step, etc. Ambiguities in which possible structural expectations are drawn from different basic types are nonsystematic.

Because systematic ISA ambiguities are predictable, it is possible to exploit special

knowledge about them and tailor special disambiguation mechanisms to them. For nonsystematic ambiguity, it appears that the best tools we have available are general heuristics and a general disambiguation mechanism. [62] The mechanism is "general" in the sense that it is not keyed to particular ISA ambiguities. In this paper, I restrict myself to systematic ISA ambiguity; the next section discusses the application of special knowledge associated with ISA types. Before that, however, it is useful to take a general look at the dependencies that must be taken into account in the disambiguation process.

10.4. Dependencies in ISA Disambiguation

A computational approach to ISA disambiguation has several inherent problems. For those utterances with only an immediate speech act force, the recognition process will not necessarily be simpler than for utterances with more than one speech act force. For example "Are you able to open the door?" could have the immediate force ASK only, with no secondary force. For this type of utterance, it will often be necessary to consider the possibility of a secondary force in order to rule it out.

Also, for utterances where an action could be implicit, it is often necessary to determine whether P1 intends an action (and what that action is) as part of the process of determining the immediate or secondary force. For example, 10.5 could have the immediate force STATE, or, say if P1 were blackmailing P2, it could have an immediate REQUEST force.

10.5 I want a million dollars.

---

62. In [3] I discuss the use of ordering of match attempts as one part of a mechanism for handling nonsystematic ambiguity in task-oriented dialogue.

It is necessary to determine the likely implicit action (in this example, that P2 give P1 the money) in order to decide whether this particular action can be REQUESTed in this environment.

On the other hand, for most (and possibly all) ISAs that *do* have a secondary force, determining the immediate force can be done independently of the secondary force. This is because, as noted in Section 7.3, most ISA classes considered here are derived by only one level of indirection. [63] We thus do not get the case illustrated in Figure 10.5 where a surface form leads to two different immediate forces with two different associated secondary forces, the latter of which must be considered in determining which of the former apply. Instead the case that obtains is Figure 10.6. The computational implication of 10.5 and 10.6 is that immediate forces can be determined from the utterance and from context, independent of any consideration of the nature of the secondary forces that might apply. In the process model, we can exploit this independence of choice of immediate force where it exists, with the normal gains in simplicity and modularity that come from decoupling processes.

63. Among the ISA forms that I am considering, the only possible exception to this is the type 3 implicit-action form. Even in this case, however, we would have to have a very specialized type of form in order for determination of immediate force to be dependent on secondary force. We would have to be dealing with what I am calling a frozen ISA form (see Section 7.4). The immediate force would not correspond directly to surface form, and there would only be a secondary, not a tertiary, speech act force. Although all of this is theoretically possible, I have not found any examples in the transcripts that I have examined.

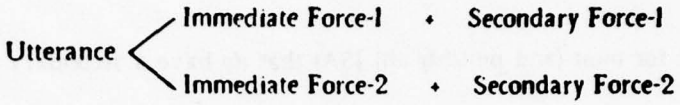
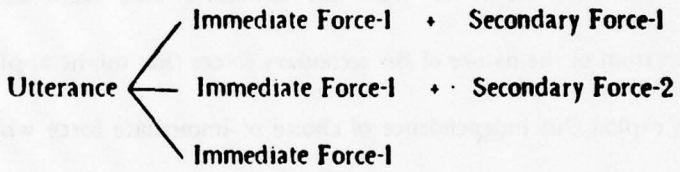


Figure 10.5. A type of ambiguity in speech act force that probably does *not* occur.



e.g.,

Utterance = "The garden is full of weeds."
Immediate Force-1 = a STATE that the garden is full of weeds
Secondary Force-1 = a REQUEST that P2 weed the garden
(e.g., in a context where this is P2's job)
Secondary Force-2 = a SUGGEST that P2 weed the garden
(e.g., in a context where P2 is deciding what to do)

Figure 10.6. The common type of ambiguity in speech act force.

11. Outline of a Mechanism for ISA Recognition

Given the framework for ISA recognition discussed in Sections 10.1 and 10.2, along with the observations about ISA disambiguation from 10.3 and 10.4, we can now examine the mechanism for ISA recognition in more detail. This section starts with a discussion of the application of the rules from Section 6. Section 11.2 considers a mechanism for recognizing frozen ISA forms, and Section 11.3 looks at the recognition of implicit-action ISAs.

11.1. Applying the Rules for Explicit-Action Forms

Recall that the ISA rules in Section 6 were based on properties of speech acts as a class, preconditions of individual speech acts, and (for some action-centered speech acts) the semantic input cases of the actions. In this subsection, I outline a mechanism for recognizing rule-based ISAs, splitting the discussion into the two directions identified in Section 10, SA -> ISA and ISA -> SA.

SA -> ISA is the easier direction to describe for these ISA rules. We start with the set of structural expectations that are to be matched top-down (Section 9.2). (Whether the match process should proceed sequentially or in parallel is an open question that will not be considered here.) Focussing on one speech act from this set of expectations, we can obtain the preconditions of the individual speech act by following the links in the method structure described in Section 4. It is also straightforward to get from an action named in an utterance to an appropriate set of semantic input case specifications (or, given ambiguity, a

group of such sets). Given these semantic structures, we can apply ISA rules 6.3 to 6.10 to yield patterns particular to the speech act. (Alternatively, these ISA patterns can be pre-generated and associated with the speech act.) The ISA patterns associated with the speech act, along with patterns from Rules 6.1 and 6.2 which are applicable to speech acts in general, can then be evaluated with respect to the current dialogue environment. One of the major effects of the evaluation process is to bind representations for the appropriate dialogue participants to semantic case variables in the pattern. The surface semantic representation of the utterance can then be matched against each of these evaluated patterns.

In the SA -> ISA direction we encounter predominantly non-systematic ambiguities among speech act forces. Therefore, the type of general disambiguation mechanism chosen determines how the successful matches are treated (or whether match attempts continue at all beyond an initial success).

The ISA -> SA direction for the general rules is more complicated. When the recognition mechanism is working top-down, the method structure and associated relationships conveniently group the representations on which the ISA patterns are based. Working bottom-up, we must contend with the systematic ambiguity discussed in the last section. This means that for almost every utterance we will have choices between speech acts conveyed, choices whether a given speech act force applies, or both. This introduces several problems.

First, at least for a serial matching mechanism, in the worst case every different speech act might have to be checked to see if the utterance in hand matches any of the

possible indirect forms for that speech act. Also, independent of the matching implementation chosen, deciding which speech act forces were intended by P1 can involve very complex analysis, particularly if action-centered speech acts are among the possible forces.

To cope with these difficulties, we can exploit two sources of information: knowledge about which speech act forms can be used to stand for others as ISAs and predictions of the type of knowledge that will be necessary to distinguish between competing speech act forces.

To use this knowledge, I propose a two-pass mechanism consisting of initial screening followed by analysis. Initial screening involves a match of an utterance against special entry patterns. These patterns are abstracted, if necessary, from patterns produced by the application of ISA rules such as those in Section 6 to appropriate semantic structures (e.g., preconditions). The point of the abstracting process is to produce entry patterns for which the match will involve only *context-independent* properties. This means that a match of an entry pattern involves only information explicit in the utterance or implicit information that can be derived independent of context. The authority relationship between P1 and P2, for example, is a property that would *not* be appealed to as part of the initial match.[64]

The abstracted entry patterns have associated with them analysis procedures to discriminate between alternative interpretations. Once an entry pattern is matched (due to the abstraction process there will be only one success), the associated discrimination procedure is executed. Discrimination procedures examine properties of the dialogue

---

64. The abstraction process produces patterns that are as specific as possible while still permitting a context-free match. Abstraction is done on the variables of the ISA patterns only (see Appendix B.)

context to determine which speech act forces were intended by P1. These procedures are written to incorporate the specialized knowledge that we have about individual systematic ambiguities.

For an example of the type of processing I have in mind, consider the following uses of a "Do you want..." form. (Other uses occur, but these four will do for an illustration.)

ASK+OFFER
one friend to another:
Do you want to take my car?

ASK+REQUEST
one roommate to another:
Do you want to answer the phone?

ASK+SUGGEST
parent to child:
Do you want to try to build an arch?

ASK
one student to another:
Do you want to go back to school?

Any of these examples pass an initial match of a representation informally conveyed by: [65]

interrogative: whether P2 wants nonverbal action
or verbal action whose destination is not P1
except for: 1. actions involving joint agency of P1 and P2
2. the action of remembering a previous REQUEST

---

65. This pattern illustrates the abstraction process. Since abstraction only happens for variables, only the representation for the action desired is affected. For OFFER, this action is "some action that complements P1's part of the action" (OFFER IV.). For REQUEST, the action is "a non-verbal action or a verbal action whose destination is not P1, except for the action of remembering a previous REQUEST." For SUGGEST, actions with joint agency of P1 and P2 are excluded, while for ASK there is no restriction. To produce the entry pattern, the restriction on the action in OFFER is ignored, because context dependent information might be required to verify it. This leaves SUGGEST and REQUEST with restricted actions. An entry pattern is then formed using both of these restrictions.

Once any of the four types of utterance above has passed the initial screening, we need a procedure to determine the correct speech act forces. To discriminate between the various speech act possibilities, I suggest the tests listed below applied as shown in Figure II.1. [66]

Tests for REQUEST:

1. Does the action named fall within some authority, complementary role, or "be helpful" obligation of P2 to P1? (Recall the discussion of these obligations in Section 3.)

2. Would one generally expect that P2 could successfully take responsibility for the action named?

The ASK Block:

1. Is there a reason why P1 would want to know the literal interpretation of the utterance without necessarily wanting to set in motion the action named?

2. Does the context preclude setting the action named in motion? This would be the case if the context is not a normal one for the action and there is no reason to assume that the standard conditions have been overridden.

Tests for OFFER:

1. Would one generally expect the primary benefit of the action named to be to P2 in this context?

2. Is P1 to be an active participant (i.e. AGENT or CO-AGENT) in the action?

3. Does it look like P1 can successfully take responsibility for an action that complements the action named (e.g., "lend a car" for "take a car")? One issue involved here is whether any objects of a transfer fall under P1's ownership, control, etc.

---

66. Recall from Section 10.3 that these tests are to be done with respect to the world model that P2 believes that P1 is using.

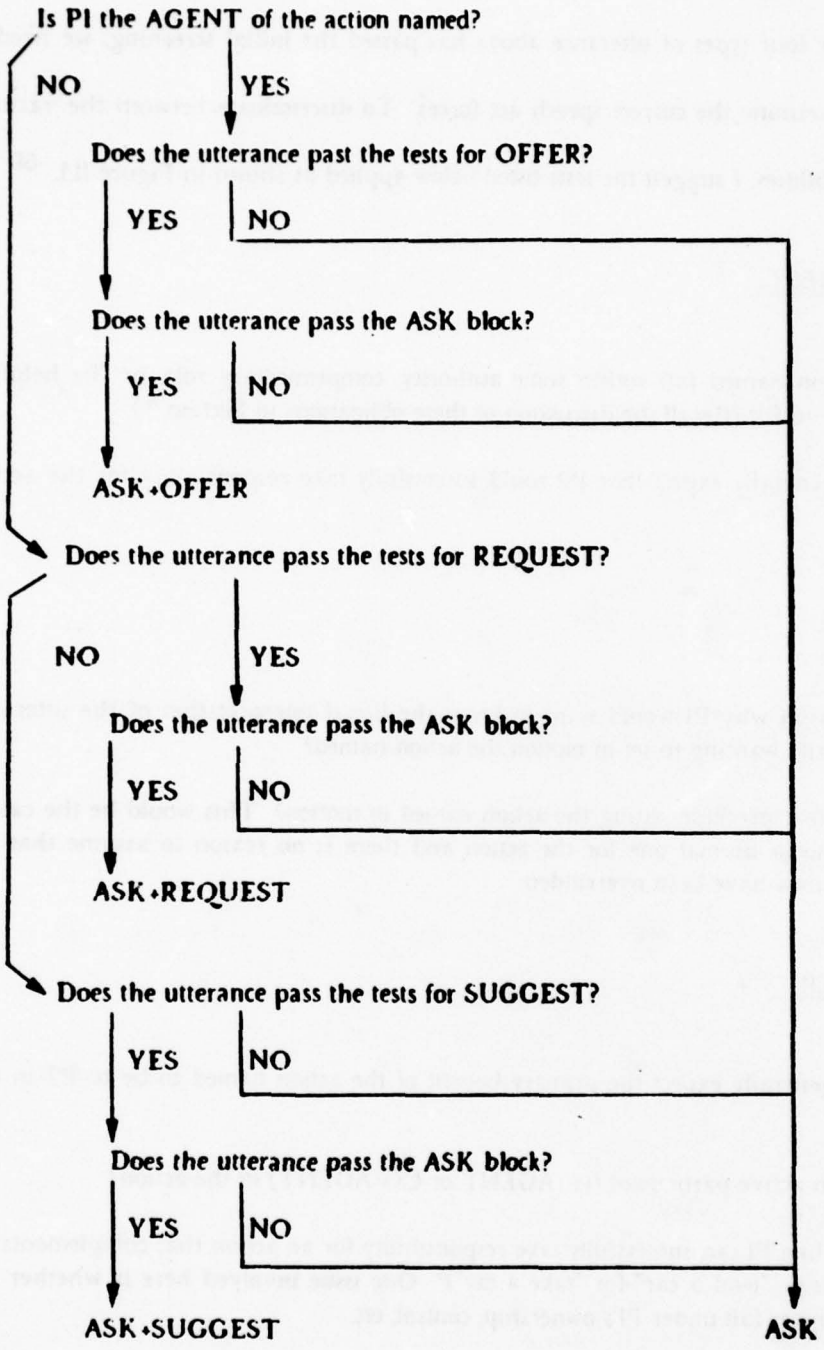


Figure II.1. General organization for discriminating between four interpretations of the form "Do you want <action>?"

Tests for SUGGEST:

1. For each *equal* authority relationship between P1 and P2, can the action named be construed as falling within the relationship?

2. Is P1 *not* to be an active participant in the action named?

3. Would one generally expect the primary benefit of the action named to be to P2 in this context?[67]

4. Would one generally expect that P2 could successfully take responsibility for the action named?

The set of questions and Figure II.1 are intended for a limited set of interpretations of "Do you want..?" forms, but they are representative of other discrimination procedures. The tests given include checks of the speech act preconditions (see Section 3), as well as checks of semantic case fillers for the speech act or action named. Some of these, e.g., the third test for SUGGEST, may be more restricted than the comparable precondition or semantic case specification. The following criteria were used to decide which preconditions to *exclude* from the tests:

1. Is this precondition shared by all of the speech act possibilities?
Preconditions shared by all alternatives may be worth verifying, but not as part of a discrimination procedure.

2. Is this precondition being questioned or stated by the ISA form?
This criterion rules out REQUEST (III), SUGGEST (VI), and OFFER (VI) for the "Do you want.." example.

3. Is the speech act usually the only evidence for this precondition?

---

67. This condition seems to be limited to the case in which the utterance is acting as an initiator (see Section 9). Utterances that convey a SUGGEST act but which occur in a context of stronger structural expectations seem to dispense with this condition and use the more general SUGGEST precondition (VII) instead. (I have suggested a basically top-down (SA --> ISA) treatment of this latter type of utterance.) It is not surprising to see stronger restrictions on the range of initiators, since the context needed for disambiguation is weaker.

> An example is the obligation in SUGGEST (IV) that the DESTINATION of the SUGGEST consider the action suggested. Such obligations rest on very general properties of cooperative activity, for which there will not usually be good particular evidence. Similarly, there is frequently no evidence other than the speech act for intentional states of another individual. Preconditions of this type may be worth checking as a tie-breaker, but the chances of getting useful information are small enough that these preconditions should probably be excluded from the routine checking of a discrimination procedure.

One general observation about the example procedure relates to the independence issues discussed in Section 10.4. Recall the point made there that immediate forces can usually be determined independent of secondary ones. This independence is reflected in Figure 11.1, where the emphasis is on determining what, if any, secondary force applies. In that case (although not in all) the fact that the immediate force was ASK fell out from the first-pass match of the entry pattern for "Do you want..." forms.

A discrimination procedure such as the one described will not always produce a single answer. Lexical ambiguities within the utterance might lead to several different alternatives. In addition, particular tests can lead to multiple alternatives, especially those that involve rich contextual structures such as role relationships. For cases such as these where a discrimination procedure fails to sufficiently narrow the possible interpretations, the recognition mechanism can appeal to other sources of information. One useful source is the implications of particular ISA choices (e.g., with respect to politeness, confidence of speaker, formality of the environment, urgency of the speech act, etc.). Another is the raw likelihood that a given form is used to convey an individual speech act or set of speech acts. The primary value of these sources would appear to be in ruling out interpretations.

The next issue that must be addressed is the generality of the discrimination

procedure associated with the four types of "Do you want..." utterances. Can the discrimination procedure be used for other types of ISA forms? In general, I think that the answer is no. Although discrimination procedures for ISA forms will often appeal to similar sorts of information and use similar tests, they differ in the following ways:

1. Deletions from and additions to the set of speech act possibilities can alter dependence relationships, which in turn affect the order in which tests are done.

2. Some forms are more commonly used for a speech act than others. Even where the same speech act set is possible for two patterns, differences in raw likelihood might affect the order in which tests are performed.

At this point it is not clear exactly how many discrimination procedures will have to be written. The number of procedures is related to the number of systematic ambiguities, and this number depends on the number of speech acts and the number of ISA forms known to a system. We are saved from the spector of infinitely proliferating discrimination procedures by the original intent of these procedures. Discrimination procedures are specially tailored to a specific systematic ambiguity. We need only write as many different discrimination procedures as there is specialized information. Where specialized information is lacking for an ambiguity, the entry pattern can have an associated list of possible speech act forces and a call to a general discrimination procedure. The core of this general procedure would be a check of all the preconditions of the possible speech act forces.

***

This has been a good deal of detail, and a summary may be helpful before going on. First, in the SA -> ISA direction the recognition mechanism starts with one or more speech

act expectations. The task is to find or derive the appropriate ISA patterns, evalate these with respect to the current dialogue context, and match the utterance against them. The method and other semantic linkages discussed in Part I provide a way to organize the ISA patterns and thus make the SA -> ISA operation relatively straightforward.

In the ISA -> SA direction, the recognition mechanism starts with the ISA pattern, and its task is to identify one or more appropriate underlying speech acts. For this, I have suggested special discrimination procedures associated with ISA patterns (or abstractions of them). These discrimination procedures capitalize on the systematic ambiguities displayed by ISAs.

II.2. Recognizing Frozen ISA Forms

Recall that frozen forms were defined in Section 7.4 as follows:

> A frozen ISA form is an ISA form with an immediate force *that does not correspond to the surface form*. For such an utterance, one or both of the following obtains:
>
> 1. It can or does display syntactic behavior atypical of the rest of the form.
>
> 2. It can or does display cooccurrence behavior atypical of the rest of the form.

To record this special behavior, frozen forms require separate patterns, distinct from patterns derived from general properties of speech acts and actions. Also, since we want to use different disambiguation schemes according to the direction of the match (a general disambiguation scheme for SA -> ISA and discrimination procedures for ISA -> SA), we need one set of frozen form patterns for each direction. While this may sound expensive, I

have found it to be useful, because other information related to the path-building process can be associated with the different sets of patterns. (The "key" representations and "way" methods in [3] constitute such a split in patterns for frozen forms.)

Besides having patterns distinct from the general ISA patterns, frozen forms will often have specially designed discrimination procedures. One reason for this is that the special surface behavior of some frozen forms will eliminate one or more possible interpretations altogether. Consider, for example, the familiar pair 5.3 and 5.4:

5.3 Can you please close the door?
5.4 Are you able to please close the door?

Without *please*, 5.4 can be either an ASK only or secondarily a REQUEST; the *please* in 5.3 eliminates the possibility that it was intended as an ASK.

Even forms that do not show special behavior themselves may be more prone to certain interpretations, justifying the use of separate discrimination procedures. For example, II.1 is more idiomatic as a REQUEST than as an ASK. To convey an immediate ASK force, a form such as II.2 is generally preferable.

II.1 Will you shut the door?
II.2 Are you going to shut the door?

(Example II.2 could also convey a secondary REQUEST force, but that is not my point here.)

Frozen ISA forms, then, require separate patterns and, in at least some cases, separate discrimination procedures. They also require special treatment in the recognition matching algorithm. Because frozen forms are special cases, it is reasonable to try to match them first

(in a serial algorithm) or to assign a higher priority to a successful frozen form match (in a parallel algorithm). This policy could result in some non-frozen utterances being treated as frozen ones, e.g., a "Can you <action>?" form intended as immediate ASK, secondary REQUEST. This simplification does not seem to me to be serious. Often, information will be too sparse for P2 to determine how seriously P1 intends a question such as the "Can you <action>?" form. Since such a wide range of responses to ISAs is possible (as noted in Section 5.2) we are not restricting the flow of the dialogue by treating all such forms as frozen.

11.3. Recognizing Implicit-Action ISAs

In this subsection, we consider the problem of integrating the recognition of implicit actions in with other aspects of the recognition of ISA forms. Discussion will concentrate on the "when" of implicit-action processing. Section 11.3.1 considers the treatment of ISAs without tertiary forces; the special problems associated with tertiary forces are discussed in Section 11.3.2.

One important topic that is not considered here is the actual search process involved in identifying a particular implicit action. To a large extent, this process is dependent on general solutions to knowledge representation problems; it is not a process unique to ISAs. Nevertheless, issues of search strongly motivated the distinction among ISA types in Section 7.5. Similarly, the tie-in between the rules in Sections 6 and 7 and the action representation in Section 4 was partly motivated by the search problem. Also relevant to this problem is the discussion of knowledge sources in Section 12; these sources are expected to guide and constrain the search.

11.3.1. Integrating Implicit Actions into the ISA Mechanism

As usual, the ISA recognition problem will be broken down according to the direction of match, and, as usual, the SA -> ISA direction is the easier one to describe. Recall that the implicit-action rules given in Section 7.5 combined a pattern intended to match the surface semantic representation of the utterance with a pattern built from a general ISA rule (Section 6). That is, implicit-action rules can be thought of as beginning where explicit-action rules leave off.

The straightforward approach to the SA --> ISA match for implicit-actions starts with the explicit-action patterns evaluated to the current dialogue environment. Implicit-action rules are applied to these evaluated patterns. In the course of this process, some implicit-action rules will be found to be inapplicable, and the others will be evaluated. The utterance can then be matched against these evaluated implicit-action patterns. Both explicit- and implicit-action patterns, evaluated with respect to the current dialogue context, could be matched in whatever order or combination desired, allowing, of course, for the dependence of implicit-action forms on explicit-action ones.

There are several possible variants to this approach to matching implicit-action patterns; the order of evaluation and rule application depends to a high degree on the general representation scheme chosen. Whatever the final choice, however, the important point about the SA --> ISA direction is that patterns evaluated with respect to the current dialogue environment can be quite specific. In many cases, a particular implicit action will already be represented in the evaluated pattern. For example, utterance 10.1 in the context discussed in Section 10 would match a SA --> ISA pattern that explicitly contained both a representation for "signature" and a representation for a signing action.

10.1 I need your signature.

Thus, for such examples there is no need to search for an implicit action. For other SA --> ISA examples, the decision about whether a particular action is intended and the search for the action can be done as part of the pattern match.

Turning now to the ISA -> SA direction we have the problem of integrating implicit actions into the discrimination procedure mechanism used for both general ISAs and frozen forms. The principal question is at what point in the process the choice of -- or search for -- the implicit action should take place. This question requires careful analysis because the process of identifying an implicit action appears to be an inherently expensive one. For most of the implicit-action types, the process is heavily dependent on a wide range of knowledge sources, where the exact source is not predictable beforehand.

Identification of implicit actions can either be done at a fixed place in the bottom-up recognition process or at different places in the process depending on the speech act possibilities involved. The obvious candidate for a fixed place is the first matching pass. That is, the implicit action would be determined on entry to the discrimination procedure as part of matching the entry pattern. The evidence, however, points against identifying implicit actions during the initial match pass. First, the implicit-action rules as written would not be easily applicable to entry patterns. This is because implicit-action rules are specific to individual speech acts, and entry patterns are not. Even if some implicit-action rules were generalized, the link between implicit actions and individual speech acts goes deep. For example, 11.3 can be either an OFFER by *P1* to give *P2* a pen or a REQUEST

that *P2* give *P1* a pen.[68]

11.3 Do you have a pen?

The choice between these two alternatives can only be resolved within the discrimination procedure. There appears to be no advantage to generating both representations early and carrying them forward.

A similar problem arises for forms where implicit action ambiguities occur for a single speech act force. For example, consider the following form:

11.4 Do you have <INSTRUMENT> for <action>?

If the AGENT of the action is not specified, then there are two different REQUESTs among the possible interpretations. The first possibility is a REQUEST that P2 give P1 the INSTRUMENT, and the second possibility is that P2 take responsibility for carrying out the action using the INSTRUMENT.

Given the close relationship between implicit action choice and speech act force, then, it appears that the choice of implicit action should occur within the appropriate discrimination procedures, under explicit control of the system builder.

11.3.2. Recognizing Utterances with Tertiary Forces

As was the case for other ISA classes, the fact that utterances can convey a tertiary

---

68. For an OFFER, example 11.3 is derived by an application of an implicit-action rule analogous to Rule 7.4 to the result of the general ISA rule 6.1 Clause iv. For REQUEST, example 11.3 comes from the application of the implicit-action rule 7.2 to the result of the application of the general ISA rule 6.4 to Precondition (11).

speech act force raises not only the problem of recognizing when such a force is intended, but also the problem of recognizing when such a force is *not* intended. Recall that among the ISA types considered here, type 3 implicit-action forms are the only ones that may systematically convey tertiary forces.

5.18 Close the door.
5.19 It's cold in here.
5.20 Must I tell you that it's cold in here?

Using our standard examples, recall that in an appropriate context 5.20 has three forces: an immediate force of ASK that corresponds to the surface form, a secondary force of a representative that conveys information like that in 5.19, and a tertiary force of REQUEST that conveys information like that in 5.18.

In the SA -> ISA direction, tertiary speech act forces present no special problems. In addition to the normal match of ISA patterns outlined in the first subsection, the Interpreter can apply the appropriate rules a second time to type 3 forms to find the possible double indirection patterns. A match of one of the patterns so produced -- with accompanying identification of an implicit action -- indicates that a tertiary force is a possibility. After a successful match of a triple-force utterance, the Interpreter builds three Event Tree paths, one for each force.

Once again, the ISA -> SA direction is more difficult. The main problem presented by a double indirection in bottom-up processing is that a decision about whether the first possible indirection has been taken (e.g., ASK to the representative in 5.20 above) can be contingent on processing related to a second indirection (e.g., a representative to

REQUEST). More specifically, the existence of an action appropriate to be an implicit action for a type 3 REQUEST form could be evidence not only for the second indirection, but also for the first. This means that we definitely cannot think of the bottom-up recognition process for double indirections as two subprocesses acting independently of each other. The decision about whether an utterance such as 5.20 is intended as a direct ASK, an ASK plus an indirect representative, or an ASK plus a representative plus a REQUEST may be computationally messy because processing cannot be subdivided in all cases.

Although double indirection clearly poses special problems, I do not want to give the impression that the situation is hopeless. Note first that the analysis phase described in the last subsection uses general procedures. The problem, then, is not finding a way to handle double indirection within this framework but, rather, finding a *good* way. It appears that the answer to most of our problems can be found in the data itself. In the transcripts that I have examined, the proportion of type 3 ISAs to other forms is small; moreover, of the type 3 forms, only a subset are based on a double indirection. More important for bottom-up processing, only a handful of different forms need to be considered as possible double indirections (whether or not the utterances actually are).

Beyond simple numbers, note that the literal interpretation of double indirections appears to require strong contextual support. Consider, for example, I1.5 and I1.6:

I1.5 Do you know that it's cold in here?
I1.6 I want you to know that it's cold in here.

Without a specialized context (e.g., a psychological experiment) it is difficult to imagine I1.5 or I1.6 uttered as an initiator with only the literal interpretation intended.

The need for special contextual support is the most important property of the data for bottom-up processing, because it means that outside of such contexts the first indirection can be assumed with relative safety. The task then becomes to determine whether or not a tertiary force applies, e.g., whether 11.5 is an ASK and a STATE or an ASK, a STATE, and a REQUEST. This task is similar to the processing of the possible single indirection "It's cold in here." So, while in theory double indirections are more difficult computationally, it appears that in practice only rare -- and recognizable -- cases require more processing power than a comparable single indirection.

12. Knowledge Sources for ISA Recognition

Throughout this paper I have emphasized the importance of context in the interpretation of ISAs. While a complete account of the role of context in ISA recognition is not possible at this point, we can begin to identify the knowledge sources involved. This section examines the following six types of knowledge:

1. domains of discourse
2. complementary role relationships
3. special properties of P1
4. special properties of P2
5. exceptional circumstances
6. the course of the interaction to date

These six knowledge sources play an important role in the interpretation of implicit-action ISA forms, e.g., in identifying the intent of "It's cold in here" uttered as a REQUEST for P2 to shut the door. Although implicit-action forms will be used as examples in this section, this is not the only place that the six knowledge sources can be used. The sources also play a role in supplying semantic case information that is left implicit in explicit-action forms. For example, the INSTRUMENT is left unspecified in example 12.1, but, if P1 uttered 12.1 while handing P2 a bread knife, then the case assignment would be clear.

12.1 Will you cut the bread?

Finally, the six knowledge sources can supply information needed to disambiguate forces of an ISA, for example the possible readings of a "Do you want ..." form discussed in Section II.

Two additional observations must be made about the list above. First, the six types of knowledge are not necessarily seen as independent. We can expect interaction between sources, especially between any of (2) through (6) and any of the other sources on the list. Second, recall that I am assuming for the sake of exposition that P1 and P2 share an identical world model with respect to this knowledge. This assumption will not be correct in all cases. Sometimes ISAs uttered by P1 will be interpreted by P2 in terms of P2's view of P1's view of the six knowledge sources.

12.1. Domains of Discourse

In this section, *domain* is meant to include:

1. the spatial environment: locations such as home, office
2. the social environment: e.g., public transportation, parties, school lectures
3. the general task environment: e.g., eating a meal, running a business [69]

In talking about the relation of domains to the interpretation of ISAs, I wish to make a case for the existence of *routine actions associated with particular domains*. By *routine* I wish to emphasize the concepts that appear in dictionary definitions, such as prescribed, regular, customary, habitual. That is, routine actions associated with particular states in particular domains are *not* actions that are arrived at by intricate creative or planning processes. They are rather humdrum actions whose indications for use are common knowledge.

---

69. I specify *general* in the task environment, because specific tasks will be discussed separately in Section 12.6. General task environments are activities that include a collection of independent tasks, i.e. tasks that are not subtasks of each other.

To illustrate, I return to "It's cold in here" used as an indirect REQUEST for a particular action. In an average home, the routine cold-weather responses to this utterance might include:

1. check if windows or doors are open or are admitting drafts
2. check to see if more sun can be let in
3. turn up heat
4. add clothes, blankets, etc. to anyone who is cold

The strict ordering of the alternative actions is not obligatory in all cases, but it is appropriate in some. Note that such routine actions can be associated with very specific domains. Another household might use the same responses but order them differently. A third household might include throwing a log on the fire as one of the actions. For this example, the domain may be not only a particular home but even a particular room. To represent the routine actions of a domain, we can group them by the goal(s) that they achieve and write choice methods for each cluster. These choice methods will contain the known tradeoffs between the actions, as well as any ordering that is appropriate.

Returning now to speech acts, one way to communicate an implicit action (or other implicit information) in an ISA is to draw on the routine actions associated with the particular domain. If a particular action is intended by PI, it must be distinguishable from among this set of possibilities. Otherwise, the ISA is not successful. Where the alternative actions have a strong a priori ordering, we can filter the alternatives according to what is possible in the domain and add criteria provided by the other knowledge sources discussed in this section. Where there is no particular order of choice, either other knowledge sources should be applicable or other routine actions should be inapplicable in the environment.

Assuming that an action was intended, if the implicit action is still ambiguous, then either P2 has too rich a set of possibilities or P1 has been vague.

12.2. Complementary Roles of P1 and P2

Another source of information for clarifying ISAs is the role, i.e., a set of actions which may be associated with a dialogue participant. In this subsection, I am considering only complementary role relationships, e.g., doctor/patient or conductor/passenger. Roles of dialogue participants that are not complementary can also affect the interpretation of ISAs; they are included under properties of P1 and P2 (Sections 12.3 and 12.4). Like domains, roles exhibit different levels of aggregation. On an assembly line, a role might be a relatively small set of actions or even a single one, such as tightening a particular bolt. The role for, say, a homemaker could contain many more actions, possibly grouped into related sets such as cooking, house cleaning, home repair, and so forth. Here, *homemaker* is what I will call a *role name*; the word *role* will be reserved for the action set.

Roles are particularly important sources of information about ISAs in task-oriented environments, where formal job descriptions often exist and where labor is often explicitly divided. They also come into play in more informal environments, where tasks and responsibilities are divided by custom, by social role, or by agreement.

An example of an implicit-action ISA based on a complementary role relationship is 12.2 when addressed to P2, an auto mechanic (a role name) at P2's place of business (a domain).

12.2 The clutch on my van is gone.

Under the conditions specified, I2.2 can count as a REQUEST that the mechanic fix or replace the auto part, in this case the clutch. This is one of the actions in an auto mechanic's role. Utterance I2.2 can also have the force of a simple STATE, for example if it is directed to a friend encountered while waiting in line. For this example in this environment, it takes the mechanic's role to make an appropriate action, and hence a REQUEST sense, clear. To ask the friend met in line for a ride, one would generally need a more specific REQUEST form in addition to I2.2.

One question that must be asked about roles is what effect multiple roles have on the interpretation of ISAs. Participants in a dialogue often share more than one complementary relationship. While domain introduces some possible roles (e.g., host or hostess at a party), other roles arise strictly from the relationship of P1 and P2. These multiple relationships do not seem to confuse the interpretation of ISAs as much as one might suspect. Consider an example in which P2 is both P1's doctor and P1's friend. If P1 describes a pain to P2 at a party, it is probable that P1 is REQUESTing more than sympathy; very likely the appeal is to P2's special expertise and the implicit action is a diagnosis. Where more than one complementary role relationship exists, the default seems to be to appeal to the role with the most powerful action in the area of the REQUEST.[70] This is not to say, however, that confusion cannot occur, and in this case, a clarifying question can be asked.

---

70. This is subject, of course, to constraints. Example I2.2 uttered to a mechanic at a party would seldom be seen as a REQUEST to replace a clutch right then and there. This reading could be ruled out by the absence of the necessary tools and parts, by the party domain (because it conflicts with standard party activities), or by some other condition. Note that the first two conditions named correspond to the second blocking question that leads to an ASK interpretation in the example worked out in Figure II.1.

12.3. Properties of P1

While complementary roles relate to the mutual relationship of P1 and P2, properties of P1 independent of the relationship can also affect the interpretation of an ISA. One important case mentioned above is the non-mutual role, such as an appeal to the medical expertise of P2 where P1 is not a patient. Another important class of properties includes the likes and dislikes of P1. Consider, for example, 12.3 uttered as a REQUEST:

12.3 I have to be in Cleveland by 9:00 tomorrow morning.

Let us assume that 12.3 is uttered in an office (a domain) to a secretary (a role name). Let us assume further that the routine response to such a REQUEST in this office is to make travel arrangements, in particular to book a flight if the distance is appropriate. Consider the case, however, in which P1 is afraid to fly and has made P2 aware of this fact. In this case, some other routine action may be understood, such as making train reservations. This default would be associated with P1.

In general, properties of P1 may block alternative actions suggested by other knowledge sources (and hence perform a disambiguation function), they may suggest totally new alternatives, or they may cause a reordering of existing alternatives.

12.4. Properties of P2

Properties of P2 that are common knowledge to P1 and P2 can also have an effect on the interpretation of an ISA.

12.4 Tabby needs a new home.

If P2 is allergic to cats, P1 knows it, and P2 knows that P1 knows, then 12.4 will not (assuming the good will and good memory of P1) be a REQUEST that P2 adopt Tabby.

In general, properties of P2 are of a type and behavior similar to properties of P1. There is, however, an added dimension. In deciding whether an interpretation is affected by properties of P2, P2 must determine both whether P1 knows about the property and whether P1 seems to be aware of it at the time of the REQUEST. To make these decisions, P2 will usually be operating with highly incomplete knowledge. If there is any question of the right choice, and if anything significant rides on the choice, P2 will often state the property (e.g., 12.5 in response to 12.4) as a verification of the choice.

12.5 If I weren't allergic to cats, I'd be happy to take Tabby.

12.5. *Exception Conditions*

The last four subsections have enumerated sources of knowledge, with the stated assumption that standard conditions prevailed. Interpretations of ISAs can be affected dramatically, however, by the existence of exception conditions. Exception conditions are not a totally separate source of knowledge but are integrated into all four categories discussed so far: domain, role, and the special properties of P1 and P2.

An example of an appeal to a domain-related exception condition is 12.6 uttered at home in a power failure.

12.6 It's dark in here.

If P1 is aware of the power failure, then it is very unlikely that 12.6 is a REQUEST that P2

flip the light switch. Due to the exception condition, in this particular domain getting a flashlight might be the preferred action (as well as the one intended by P1).

Exception conditions are not limited to domains. An exception condition in a role might arise when a piece of machinery is broken and a different manufacturing procedure must be used. Finally, for P1 and P2, an exception condition might be a broken leg.

Exception conditions not only block or reorder standard alternatives; they may also bring totally new sets of alternatives of their own. For example, the use of a flashlight, while *a* solution to darkness in a standard situation, is not a *routine* solution in most domains. Note that an implicit action, or other aspect of an ISA, can only be "clear from context" in an exception situation when the exception condition has its *own* routine actions associated with it. We are still talking, then, about *non-creative, routine* actions. The difference is that these actions are associated with non-routine conditions.

12.6. The Course of the Interaction to Date

This brings us to the last category, the effect of the interaction -- both linguistic and non-linguistic -- between P1 and P2. Several different things may occur here. First, information may be gained in the course of the dialogue that directly pertains to the categories already discussed. P1 may explicitly specify the domain, e.g. "Let's view this problem geometrically." One of P1's roles may also be specified explicitly, e.g., "Speaking as the owner of this property, I want you to know that you're trespassing." Similarly, P1- or P2-based properties may be discussed explicitly in the dialogue, or an exception condition may be noted.

Interpretations of ISAs may also be affected by specific tasks that are now or were previously underway in the course of the current interaction, as well as the goal structures of P1 and P2 revealed in the dialogue. [71] An example of the effect of specific task is the interpretation of example 12.7.

12.7 We have twenty cartons of paper clips.

If the domain is an office and the specific task is taking inventory, then 12.7 can be interpreted as a simple representative. If the task is rearranging the supplies, then 12.7 lends itself naturally to interpretation as a REQUEST for advice on where to put things or for help in moving them. An example of the effect of discussion about goals is 12.8, which can be seen as part of an extended exchange related to planning the day.

12.8
12.8.1 P1: I think I'll go to the A&P today.
12.8.2 P2: We're out of flour.

It is reasonable to construe P2's utterance in 12.8 as a REQUEST that P1 buy flour on the planned shopping trip. To see that this interpretation of 12.8.2 is dependent on 12.8.1, compare the following:

12.9
12.9.1 P1: Look at the snow. Let's bake a cake.
12.9.2 P2: We're out of flour.

The most obvious interpretation of 12.9.2 is as a representative acting as a rejection of the suggestion in 12.9.1.

---

71. I am assuming that goals and tasks from previous dialogues and other shared experiences would already be assimilated and would be accessible as properties of the participants.

Another way that the interaction to date may affect the interpretation of ISAs is in discussion of the preconditions of actions. Consider, for example, 12.10.

12.10
12.10.1 P1: Do you play dominoes?
12.10.2 P1: I need a partner for the mixed doubles tournament on Friday.

The presence of 12.10.1 clarifies the interpretation of the particular action in 12.10.2: P1 wants P2 to agree to be P1's partner in the tournament. Without this preliminary remark, the particular action would be considerably less clear. Utterance 12.10.2 might just as well be a REQUEST for advice or for sympathy.

Although I am not attempting to describe *how* the dialogue affects the interpretation of ISAs, a case could be made for viewing the examples in this subsection as derivable from methods (Section 4) and the Event Tree and recognition framework (Section 9). This is a natural approach, because the Event Tree is one record of the linguistic and non-linguistic interaction between P1 and P2. Moreover, methods can provide strong connections between utterances and strong expectations about conditions in the environment. We can therefore expect methods and the Event Tree to be useful in modelling many (although not all) of the interactions between P1 and P2 that can have an effect on ISA interpretation.

❧

For P1 to construct a well-formed indirect speech act, there must be a reasonable chance that P2 will interpret the utterance as intended. In general, it is only by an appeal to shared knowledge that a correct interpretation becomes possible. The six knowledge sources discussed in this section give us a start toward a catalogue of the information

common to dialogue participants (or, more precisely, the information that one reasonably expects will be common). The knowledge can come out of shared experience, out of communication between participants, or from the standards individuals and institutions impose to make certain actions and knowledge "routine".

13. In Conclusion

In this paper I have presented an account of a large number of ISA forms within task-oriented dialogue. While the account was strongly motivated by computational considerations, non-computational arguments were used to justify many of the distinctions made. The major characteristics of the approach that I have taken are as follows:

-- Starting from the observation of traditional speech act theory that language is action, I have advocated a uniform representational scheme for speech acts and non-linguistic actions. The foundation of the representation is the OWL-1 method.

-- Speech acts and other actions were combined into larger patterns of action, also represented by OWL-1 methods. These larger patterns were used as a partial model of the structure of dialogue. This gave a framework in which to discuss issues of ISA recognition.

-- The central processing assumption has been that the phenomenon of indirect speech acts is too complex to admit to a single, uniform computational treatment. It is necessary, then, to identify classes of indirect speech acts that share similar computational properties and use different representations and processing strategies for the different classes.

-- I have claimed that a process model of ISA usage needs at least the following:

1. rules based on general properties of speech acts and other actions

2. patterns that represent special syntactic and cooccurrence behavior for frozen ISA forms

3. rules that embody relationships that can be used in referring to implicit actions

There is much more that needs to be done before a full process model of ISAs is achieved. First, there are the areas that were explicitly excluded, among them, hedged performatives and "second order" ISAs, implications of ISA choices, and the generation of

PRECEDING PAGE BLANK-NOT FILMED

ISA forms. Second, more in-depth analysis is needed for individual speech acts. Some of these that are particularly interesting are the speech acts for stating facts, beliefs, and opinions. Despite a great deal of attention from philosophers and some linguists, there is still much that can be done in this area. Third, notions of context need to be more fully worked out. A fourth major area is reasoning for disambiguation; we are far from a complete understanding of how to block competing, but incorrect, interpretations of ISAs.

Finally, I would like to comment on the implications of all of this for human-machine dialogue. Judging from the open problems I have listed, one might expect that computer systems that reliably recognize ISAs would be many decades away. I think, however, that a partial solution to the problem is much more accessible to us. Using the speech act categorization developed here, it is possible to identify useful subsets of ISAs. This already is an advantage, since it is often very difficult to identify useful subsets of English; frequently implementers find themselves in what amounts to an "all or nothing" situation. Moreover, given the categorization, it will be easy to describe what a system actually *does* for ISAs. Since the categories were justified to a large extent by surface English behavior, they should appear natural to speakers of English. Again, this is an area in which implementers have had only mixed success.

Another way in which the implementer's job can be greatly simplified is by carefully restricting the speech acts accepted and the tasks available for discussion. (Both restrictions, of course, must be made clear to the user.) Limiting the types of speech act accepted means that some secondary forces are simply not possible in the environment. With only a few tasks handled, it is not as difficult to identify an implicit action; fewer tasks mean that the task intended will usually be "clear from context."

Thus, while many problems remain to be solved, I feel that not only is a computational theory of ISAs well on its way, but we are at the point where ISA processing can be done cleanly, if not yet completely, in computer dialogue systems.

References

[1] Allen, J. A plan-based approach to speech act recognition, Technical Report No. 131/79 (February 1979), Department of Computer Science, University of Toronto, Toronto, Canada.

[2] Austin, J.L. *How to Do Things with Words*, Oxford University Press, New York, 1962.

[3] Brown, G.P. A framework for processing dialogue, LCS TR-182 (June 1977), Laboratory for Computer Science, M.I.T., Cambridge, MA.

[4] Brown, G.P. Linguistic and situational context in a model of task-oriented dialogue, to appear in L. Vanya and K.J.J. Hintikka (eds.): *Models of Dialogue*, D. Reidel, Dortrecht, Holland.

[5] Bruce, B.C. Belief systems and language understanding, BBN Report No. 2973 (January 1975), Bolt, Beranek and Newman, Inc., Cambridge, MA.

[6] Cohen, P.R. On knowing what to say: planning speech acts, Technical Report No. 118 (January 1978), Department of Computer Science, University of Toronto, Toronto, Canada.

[7] Davison, A. Indirect speech acts and what to do with them, in: Cole and Morgan (Eds.) *Syntax and Semantics*, vol 3., Academic Press, New York, 1975.

[8] Deutsch, B.G. The structure of task oriented dialogs, *Proceedings IEEE Speech Symposium*, Carnegie-Mellon University, Pittsburgh, PA (April 1974).

[9] Fraser, B. Hedged performatives, in: Cole and Morgan (Eds.) *Syntax and Semantics*, vol 3., Academic Press, New York, 1975.

[10] Gordon, D. and Lakoff, G. Conversational postulates, in: Cole and Morgan (Eds.) *Syntax and Semantics*, vol 3., Academic Press, New York, 1975.

[11] Grice, H.P. Logic and conversation, in: Cole and Morgan (Eds.) *Syntax and Semantics*, vol 3., Academic Press, New York, 1975.

[12] Grosz, B. The representation and use of focus in a system for understanding dialogs, *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, MA (August 1977).

[13] Hawkinson, L. The representation of concepts in OWL, *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence*, Tbilisi, Georgia, USSR (September 1975).

PRECEDING PAGE BLANK-NOT FILMED

[14] Labov, W. and Fanshel, D. *Therapeutic Discourse*, Academic Press, New York, 1977.

[15] Lakoff, R. The logic of politeness; or, Minding your p's and q's, *Papers from the ninth regional meeting of the Chicago Linguistic Society*. University of Chicago Department of Linguistics, Chicago (1973).

[16] Levin, J. and Moore, J. Dialogue games: Meta-communication structures for natural language interaction, *Cognitive Science* Vol.I, No. 4, October 1977.

[17] Long, W.J A program writer, LCS TR-187 (November 1977), Laboratory for Computer Science, M.I.T., Cambridge, MA.

[18] Martin, W.A. OWL: A system for building expert problem solving systems involving verbal reasoning, 6.871 Course Notes, unpublished, M.I.T., Cambridge, MA (1974).

[19] Minsky, M. A framework for representing knowledge, in: Winston (Ed.), *Visual Information Processing*, MIT Press, Cambridge, MA 1975.

[20] Moore, J.A., Levin, J.A., and Mann, W.C. A goal-oriented model of human dialogue, *American Journal of Computational Linguistics*, Microfiche 67 (1977).

[21] Morgan, J.L. Two types of convention in indirect speech acts, in: Cole (Ed.) *Syntax and Semantics*, vol. 9, Academic Press, New York, 1978.

[22] Perrault, C.R., Allen, J. and Cohen, P. Speech acts as a basis for understanding dialogue coherence, *Theoretical Issues in Natural Language Processing-2*, University of Illinois at Urbana-Champaign, Urbana, IL (July 1978).

[23] Rubin, A. A framework for comparing language experiences, *Theoretical Issues in Natural Language Processing-2*, University of Illinois at Urbana-Champaign, Urbana, IL (July 1978).

[24] Sadock, J.M. *Toward a Linguistic Theory of Speech Acts*, Academic Press, New York, 1974.

[25] Sadock, J.M. On testing for conversational implicature, in: Cole (Ed.) *Syntax and Semantics*, vol. 9, Academic Press, New York, 1978.

[26] Sag, I. and Liberman, M. The intonational disambiguation of indirect speech acts, *Papers from the ninth regional meeting of the Chicago Linguistic Society*. University of Chicago Department of Linguistics, Chicago (1973).

[27] Schank, R. and Abelson, R. *Scripts, Plans, Goals, and Understanding*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1977.

[28] Schegloff, E. and Sacks, H. Opening up closings, in: Turner (Ed.) ***Ethnomethodology***, Penguin Books, Middlesex, England, 1974.

[29] Searle, J.R. *Speech Acts*, University Press, Cambridge, 1969.

[30] Searle, J.R. Indirect speech acts, in: Cole and Morgan (Eds.) ***Syntax and Semantics***, vol 3., Academic Press, New York, 1975.

[31] Searle, J.R. A taxonomy of illocutionary acts, in: Gunderson (Ed.), ***Language, Mind, and Knowledge***, University of Minnesota Press, 1976.

[32] Swartout, W.R. A digitalis therapy advisor with explanations, ***Proceedings of the Fifth International Joint Conference on Artificial Intelligence***, Cambridge, MA (August 1977).

[33] Szolovits, P., Hawkinson, L., and Martin, W.A. An overview of OWL, a language for knowledge representation, *Proceedings of the Workshop on Natural Language for Interaction with Data Bases*, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[34] Verschueren, J. Speech act theory: A provisional bibliography with a terminological guide, (May 1976) Indiana University Linguistics Club, Bloomington, IN.

[35] Verschueren, J. The analysis of speech act verbs: Theoretical preliminaries, (August 1977) Indiana University Linguistics Club, Bloomington, IN.

[36] Zwicky, A.M. and Sadock, J.M. Ambiguity tests and how to fail them, in: Kimball (Ed.) *Syntax and Semantics*, vol. 4, Academic Press, New York, 1975.

Appendix A. Core Dialogue Methods

This appendix contains a list of some of the core dialogue methods associated with the speech acts from Section 3. Along with the list of core methods is a description of the OBJECT case specification as well as a description of the major standard path steps. The actual representation of these methods is done in OWL-1.

ASK-AND-ANSWER
  OBJECT: a what-, where-, who-, whether-, or when-question to be answered
  MAJOR STANDARD PATH STEPS:
    P1 ASKs the question.
    P2 evaluates the query to decide whether to comply.
    P2 finds the answer.
    P2 gives the answer.

ASK-AND-DESCRIBE
  OBJECT: a thing or how-question[72]
  MAJOR STANDARD PATH STEPS:
    P1 ASKs what the description is.
    P2 decides whether to comply.
    P2 gives the description.

ASK-AND-EXPLAIN
  OBJECT: an action, a how-question, or a why-question
  MAJOR STANDARD PATH STEPS:
    P1 ASKs what the explanation is.
    P2 decides whether to comply.
    P2 gives the explanation.

---

72. How-questions are split between ask-and-describe and ask-and-explain depending on the type of information that seems appropriate. Of course ask-and-describe and ask-and-explain can also be triggered by a direct request for a description or explanation, respectively. The motivation for distinguishing ask-and-describe and ask-and-explain from ask-and-answer is that the first two will tend to be involved with longer answers that require more selection and organization of the information. Ask-and-describe and ask-and-explain are distinguished from each other by the aspects of the topic that are considered relevant in answering a how-question; for ask-and-explain, the emphasis is on causal relationships.

PRECEDING PAGE BLANK-NOT FILMED

REQUEST-AND-HELP
OBJECT: an action that can be the object of REQUEST
exception: a repetitive action
MAJOR STANDARD PATH STEPS:
P1 REQUESTs P2 to take responsibility for helping P1 with an action.
P2 decides whether to comply.
in either order:
P2 ACKNOWLEDGEs the REQUEST.
If not already understood, roles in the action are assigned to the participants.
then:
The action is carried out.

SUGGEST-AND-ACCEPT[73]
OBJECT: an action
MAJOR STANDARD PATH STEPS:
P1 SUGGESTs the action.
P2 decides whether to comply.
P2 ACCEPTs the suggestion.

---

73. Note that the action suggested is not represented as a step in SUGGEST-AND-ACCEPT. This is because I see suggestions as directed toward acceptance of an idea, not directly toward the accomplishment of actions (hence the possible response, "That's a good idea"). There is, however, a close relation between a suggestion to do an action and the action actually being carried out. To represent this, we can say that there is a general social directive that people should behave reasonably. If there is good reason to carry out an action (and no overriding objection to it) then the action should be carried out. This information would be incorporated into processes for selecting goals.

Appendix B. Matching Semantic Representations against Rules

Although the ISA rules in this paper have been presented informally, I am assuming a highly restricted relationship between the patterns specified by the ISA rules and the semantic representations of particular English utterances. This appendix summarizes the relationship.

First, we must consider the level of representation. The dialogue model I am assuming assigns more than one semantic representation to an utterance. (These representations are summarized in Appendix E.) I wish, however, to make my remarks on matching as independent as possible of the particular choice of representation level for rules and utterances. For the purpose of explanation, it is assumed that the two representations are at the same level or close enough so that the correspondences between elements can be easily identified.

This brings us to the constraints on matching. In saying that a semantic representation S matches a pattern in an ISA rule, I mean that the elements of S correspond to the elements of the pattern by any one of the following relationships:

1. An element of S matches an element of the pattern exactly.

2. An element of S has a subclass relationship to the element of the pattern. (I am assuming that the "class" relationship in "subclass" and "superclass" below is OWL-I specialization. See Hawkinson [13]. The important point here is that the two elements are related to each other by an explicit classification link in the knowledge base.)

3. An element of S has a superclass relationship to the element of the pattern.

4. An element of S may match a variable in the ISA pattern either by relationships (1)-(3) or by being an instance of the variable type. In this paper, variables are signalled by the use of *some*, *a*, or *the* in the statement of rules and preconditions. *The*

indicates the case where a variable must be matched by an instance. *A* and *some* are used where either an instance match may occur or relationships (1)-(3) may be appealed to.

5. Phrases in rules and speech act preconditions appearing in this paper in parentheses are not expected to have corresponding elements in utterance representations.

6. One of the relationships (2), (3), or (4) apply, and added constraints on what may match the pattern are either given explicitly in the pattern or supplied by the other parts of the match that have been completed.

The third relationship above is necessary, but it is significantly less common than the first two. Since parts of the utterance are more general than the pattern, in these cases context plays a role in determining that the more specific element of the pattern is being referred to.

To illustrate the fourth matching relationship, consider the following examples:

B.1 *Something good is bound to come of it* if you buy a dollar ticket.
B.2 *You'll have a good chance of winning* if you buy a dollar ticket.

Examples B.1 and B.2 can be interpreted as REQUESTs derived from Rule 6.2, repeated here.

PI can convey a speech act indirectly by --

-- a representative with the propositional content that some desirable result or results can be expected for some intended effect of the speech act.

Under this interpretation, the parts of the example in italics would correspond to the phrase "some desirable result or results" in the rule. The italicized clause in B.1 would be related to the rule by one of the relationships (1)-(3) above, depending on what stand one took on the relationship between the good and the desirable. The italicized clause in B.2 would be

considered as conveying an instance of the phrase in the rule, since winning is one possible desirable result.

In this account of matching, there is still a missing piece: the actual relationships between particular representations. In keeping with the informality of the rules that I am presenting, I do not attempt a taxonomy of representations (i.e. hierarchic structures from a world model). [74] In addition, I make no attempt to characterize the features of English that make some realisations of a rule more idiomatic than others. My purpose in discussing matching is to show that, despite the informality of the presentation, I am expecting the matching process to be a relatively straightforward, highly constrained operation.

74. Two expressions that come up repeatedly in the statement of the rules, however, are worth noting. I assume that *believe* has a superclass relationship to all idea-holding concepts, e.g., thinking, knowing, assuming, hypothesizing, etc. *Want* has a superclass relationship to all goal-holding concepts, e.g., desiring, needing, etc.

Appendix C. ISA Examples for Rules 6.1 to 6.9

The rules restated:

PI can convey a speech act indirectly by --

Rule 6.1

(i) ASKing whether the intended speech act is necessary

(ii) ASKing whether an equivalent speech act (i.e., one with the same principal intended effect) has already been performed

(iii) ASKing whether the principal intended effect can be expected to occur without the speech act

(iv) ASKing whether the principal intended effect of the speech act has already occurred

Rule 6.2

-- a representative with the propositional content that some desirable result or results can be expected for some intended effect of the speech act.

Rule 6.3 (for PI-based preconditions):

-- a representative of the topic of a PI-based precondition of the speech act.

Rule 6.4 (for P2-based preconditions):

-- an ASK of the topic of a P2-based precondition of the speech act.

Rule 6.5 (for unmarked preconditions):

-- an ASK of the topic of an unmarked precondition of the speech act.
This rule applies in a context where PI believes P2 has better knowledge of the condition in the precondition topic.

Rule 6.6 (for unmarked preconditions):

-- a representative of the topic of an unmarked precondition of the speech act.
This rule applies in a context where PI believes PI has better knowledge of the condition in the precondition topic.

PRECEDING PAGE BLANK-NOT FILMED

Rule 6.7 (for unmarked preconditions):

-- a representative of an unmarked precondition of the speech act where P1 believes he or she has better knowledge of the condition.

Rule 6.8 (for groups of preconditions):

-- a REQUEST form of an action that is a goal of P1 (i.e., A for "P1 wants A"). This rule is applicable only when the speech act has preconditions that are exact matches or specializations of the four preconditions of REQUEST.

Rule 6.9

-- an ASK about whether P2 will take responsibility for carrying out an "active" action that is a goal of P1 (i.e., A for "P1 wants A").
This rule is applicable only when the speech act has preconditions that are exact matches or specializations of the four preconditions of REQUEST.

ASK

propositional content: some question

simple form: <interrogative> e.g., Where's the mustard?

principal intended effect: that P2 tell P1 the answer to the question in the propositional content

Rule 6.1

-i- Do I need to ask you where the mustard is?
-ii- Did I ask you where the mustard is?
-iii- Do you intend to tell me where the mustard is?
-iv- Did you tell me where the mustard is?

Rule 6.2

-- I'll be able to finish these sandwiches if you tell me where the mustard is.

Precondition-based Examples

I. P1 wants to know the answer to the question.

-R6.3- I want to know where the mustard is.

II. P1 believes that P2 can tell the answer to the question.

-R6.5- Can you tell me where the mustard is?
-R6.6- You can tell me where the mustard is.
-R6.7- I believe you can tell me where the mustard is.

III. P1 believes that P2 is willing to tell the answer to the question.

-R6.4- Would you be willing to tell me where the mustard is?
-R6.4- Do you want to tell me where the mustard is?

IV. P1 wants P2 to tell P1 the answer to the question.

-R6.3- I'd like you to tell me where the mustard is.

V. P1 believes that P2 has some obligation (a role obligation, authority obligation, or general obligation to be helpful) to P1 to tell P1 the answer to the question.

-R6.5- GAP[75]

-R6.6- You ought to tell me where the mustard is.

-R6.6- It's your obligation as a member of this household to tell me where the mustard is.

-R6.7- I think you should tell me where the mustard is.

II.-V. together:

-R6.8- Tell me where the mustard is.

-R6.9- Will you tell me where the mustard is?

---

75. Although it is possible to construct contexts in which this form can be used, it seems to be only marginal. Forms such as "Shouldn't you tell me ...?" and "Don't you think you should tell me...?" are far more common. (See Section 6.3.1 for discussion of these latter forms.) Other forms derived from application of Rule 6.3 to ASK V. do not convey an ASK of the propositional content at all; for example, "Must you tell me ...?" is a REQUEST that P2 *refrain* from telling. Perhaps Rule 6.5 does not apply to ASK V. because P1 is assumed to know about any obligations that P2 has to P1, so that there is no reason to question P2.

OFFER

propositional content: some action that PI believes will be of benefit to P2

simple form: none

principal intended effect: that P2 accept PI's commitment to take responsibility for the action in the propositional content

Rule 6.1

-i- ?Do I need to offer you a ride to the airport?[76]
-ii- Has anybody offered you a ride to the airport?
-iii- GAP[77]
-iv- Have you accepted a ride to the airport?

Rule 6.2

-- I'd feel a lot better if you'd accept a ride to the airport.

Precondition-based Examples

I. PI wants to take responsibility for the action.

-R6.3- I want to drive you to the airport.

II. PI believes that PI can take responsibility for PI's part of the action.

-R6.5- Can I drive you to the airport?
-R6.6- I can drive you to the airport.
-R6.7- I assume I can drive you to the airport.

III. PI is willing to take responsibility for PI's part of the action.

-R6.3- I'm more than willing to drive you to the airport.
-R6.3- I'd be glad to drive you to the airport.

IV. PI wants P2 to perform some action that complements PI's part of the action.

-R6.3- I want you to accept a ride to the airport.

---

76. This form seems to be marginal due to the conflict between its angry connotations and the level of politeness involved in an OFFER.

77. This gap is explained by the fact that the principal intended effect for OFFER can be brought about only by the speech act; there is no independent means of achieving it. Thus, Clause iii of Rule 6.1 can never hold.

V. PI believes P2 can perform some action that complements PI's part of the action.

-R6.5- Can you accept a ride to the airport?
-R6.6- You can accept a ride to the airport.
-R6.7- I assume you can accept a ride to the airport.

VI. PI believes that P2 would be willing to perform some action that complements PI's part of the action.

-R6.4- Would you be willing to accept a ride to the airport?

VII. PI believes that P2 has an obligation (to "be helpful") to PI perform some action that complements PI's part of the action.

-R6.5- GAP
-R6.6- You must accept a ride to the airport.
-R6.7- GAP

VIII. PI believes that P2 has an obligation to P2 (by virtue of P2's own self-interest) to perform some action that complements PI's part of the action.

-R6.5- GAP[78]
-R6.6- That suitcase is heavy. You should let me drive you to the airport.
-R6.7- That suitcase is heavy. I think you should let me drive you to the airport.

IV.-VII. together:

-R6.8- Please accept a ride to the airport.
-R6.9- Will you accept a ride to the airport?

78. See discussion for ASK V.

SUGGEST

propositional content: an action
except for: actions in which P1 and P2 share common agency

simple forms:

1. What about <action>? e.g., What about joining the Marines?
2. How about <action>? e.g., How about joining the Marines?

principal intended effect: that P2 consider taking responsibility for the action in the propositional content.

Rule 6.1

-i- Need I suggest that you join the Marines?
-ii- Has anyone suggested that you join the Marines?
-iii- Are you thinking about joining the Marines?
-iv- Have you considered joining the Marines?

Rule 6.2

-- I'd be pleased if you'd consider joining the Marines.

Precondition-based Examples

I. P1 wants P2 to consider taking responsibility for the action.

-R6.3- I want you to think about joining the Marines.

II. P1 believes that P2 can consider taking responsibility for the action.

-R6.5- Could you think about joining the Marines?
-R6.6- You could think about joining the Marines.
-R6.7- I think you could consider joining the Marines.

III. P1 believes that P2 is willing to consider taking responsibility for the action.

-R6.4- Are you willing to consider joining the Marines?
-R6.4- Do you want to think about joining the Marines?

IV. P1 believes that P2 has an obligation (to "be helpful") to P1 to consider the action.

-R6.5- GAP[79]
-R6.6- You should think about joining the Marines.

---

79. See discussion for ASK V.

-R6.7- I think you should consider joining the Marines.

V. PI believes that P2 can take responsibility for the action.

-R6.5- Can you join the Marines?
-R6.6- You can join the Marines.
-R6.7- I assume you can join the Marines.

VI. PI believes that P2 is willing to take responsibility for the action.

-R6.4- Are you willing to join the Marines?
-R6.4- Would you be willing to join the Marines?

VII. PI believes that there are some reasons why the action is desirable.

-R6.5- Would it be good for you to join the Marines?
-R6.6- It would be good for you to join the Marines.
-R6.6- You'd be a credit to your sorority if you joined the Marines.
-R6.7- I believe it would be good for you to join the Marines.

VIII. PI believes that P2 has an obligation to P2 (by virtue of P2's own self-interest) to consider taking responsibility for the action.

-R6.5- GAP[80]
-R6.6- You need a new experience. You should join the Marines.
-R6.7- You need a new experience. I think you should join the Marines.

I.-IV. together:

-R6.8- Think about joining the Marines.
-R6.9- Will you consider joining the Marines?

80. See discussion for ASK V.

Appendix D: Relating Rule 7.6 to OWL-1 Methods

The notion of a basis for action defined in Section 7.5.3 can be related to the general structure of actions discussed in Section 4. By doing so, we get an additional level of detail for the definition of "basis for action" and hence a sounder understanding of Rule 7.6 and of type 3 implicit-action forms. Moreover, it should be easy to see that the additional level of detail has strong relevance for a computational treatment of these forms. I start with a look at undesirable states and then go on to look at signals and symptoms.

Undesirable states that refer to ongoing methods (i.e. actions in progress) correspond to method failure conditions. For example, D.1, which reports a failure in an ongoing copying process may be used to REQUEST that P2 fix the copier.

D.1 The paper jammed.

Other undesirable states (e.g., "It's cold in here.") will match the negation of the PRINCIPAL-RESULT or other results of some method known to P2. (We expect the method(s) to be known to P2 due to P1's obligation to avoid ambiguity; see Section 6.1.) In addition, I suspect that very frequently undesirable states will correspond to an initial condition[81] in a special planning method. For example, there might be a special planning method that contained information on "what to do when the room is cold," including alternative actions and the tradeoffs between them. (For further discussion of these methods, see Section 12.)

---

81. This could possibly be represented in a semantic input case called INITIAL-CONDITION which would be defined as a specialization of SOURCE, i.e., the point started from.

This takes care of the relationship of undesirable states to OWL-1 methods. The handling of signals and symptoms differs according to whether or not they are related to methods already underway in the dialogue encounter. Signals and symptoms related to ongoing methods will appear in a conditional step. For example, a cake-baking method that uses an oven timer will have a conditional step corresponding to the following:

When the buzzer of the oven timer goes off, check to see if the cake is done.

That is, given the signal, perform an action. Similarly, whether or not a cake method uses a signal, the following sort of conditional step might appear:

When the cake starts to smell, check to see if it is done.

Again the same type of pattern: given a symptom, perform an action. In addition to the simple presence of the states and actions in the method, we might want to label them explicitly as symptoms or signals. Note that the conditional steps referred to in type 3 implicit-actions may appear not only in the most general ongoing methods but also in substeps of them (which are themselves, of course, ongoing methods). Thus, not only the steps in the bake-cake method might be used as a basis for type 3 ISAs but also steps in the mix-batter or grease-pan methods.

For signals and symptoms unrelated to ongoing methods, it appears that one needs an additional set of representations in the knowledge base. These would contain information such as that for the following example signal:

The fire alarm is a signal of a fire or a drill.
When it goes off, leave the building immediately.

i.e <signal> is a signal of <event or process>.
When <signal> happens <carry out an action>.

A comparable example for symptoms is:

PI shivering is a symptom of PI's being cold.

i.e <symptom> is a symptom of <state>

Note that symptoms, especially, are not necessarily one-to-one. That is, shivering may also be a symptom of other states, e.g., fright. Second, I am expecting the fact that it is undesirable for a person to be cold to be independently determinable from other information in the knowledge base.

***

In summary, to relate type 3 implicit-action forms to OWL-I methods, we have called on a number of different structures. Failure conditions, PRINCIPAL-RESULTs and other results, and states in conditional steps were used to represent the undesirable states, signals, and symptoms from Rule 7.6. In addition, independent structures were suggested for some uses of signals and symptoms.

Appendix E: A Summary of Representation Levels in the General Recognition Model

The system configuration discussed in Section 9.1 presupposes various types of semantic representations. In this appendix I collect the different representations and show briefly how they relate to each other. This list is suggested as the minimal set of representations for a general recognition process.

Surface Semantic Representation (SSR)

This representation output by the parser closely follows the surface English utterance. Its salient features are:

-- Noun group references not needed for the parsing process may remain unresolved (e.g., "I saw *him*.").

-- Systematically ambiguous relationships between constituents may be unresolved if not needed for parsing. For example, "He was hit *by the door*." Here, *by* is systematically ambiguous between the agent or cause of the action in a passive construction (i.e. the door made contact with the person) and a location marker (i.e. the door is viewed as a place). The fact that the locative is less likely as an interpretation does not mean that it can be ignored as a possible reading.

-- Choices between word senses need not be made unless, again, they are forced by the parsing process. For example, in "I get it" *get* may be synonymous with *understand*, or, if *it* refers to a journal, *get* could be synonymous with *receive* or even *subscribe to*. The SSR, however, would contain only a concept corresponding to *get* unless further specialization were necessary for a completed parse.

-- ISA forms are preserved. That is, "Can you open the window?" would always have a SSR that recorded its interrogative nature and that contained a concept corresponding to *can*. Representation of any REQUEST intentions would not occur at this level.

PRECEDING PAGE BLANK-NOT FILMED

AD-A077 065 MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTE--ETC F/G 5/7
TOWARD A COMPUTATIONAL THEORY OF INDIRECT SPEECH ACTS.(U)
OCT 79 G P BROWN N00014-75-C-0661
UNCLASSIFIED MIT/LCS/TR-223 NL

3 OF 3
AD A 077065



END
DATE FILMED
8-81
DTIC

Interpreter Level Representation (ILR)

This is the main level of representation used by the Interpreter; methods are among the structures represented at this level. ILR has the following characteristics:

-- It is a case-frame oriented representation. The core set of cases are described in Section 4.

-- Among its constituent concepts ILR may contain variables. The use of these is discussed for the next representation type.

Evaluated ILR

This is an ILR with the appropriate other pieces of ILR bound to its variables. Evaluation allows general pieces of ILR to be related to current environments. Evaluated ILR is output by the module Evaluate and is used, among other places, in the structural expectations (see Section 9.2).

For example, if an order-fast-food procedure step were evaluated for three Big Burgers and a person called Julia, then the concept for Julia would be bound to the variable for the AGENT of the call, and the concept for the Big Burgers would be bound to the variable in the OBJECT specification. We thus get a transformation from one representation to another with roughly the following information content:

R1: AGENT (a person) orders fast food from the CO-AGENT (a person)
R2: Julia orders three Big Burgers from Counter-person-1

The evaluated call representation of R2 can be used to search for an appropriate method to carry out the procedure.

Transitional Representation

Transitional representation is the level of representation that appears in special matching patterns, some of which are discussed in Section II.[82] This representation is called "transitional" because it bridges the gap between SSR and ILR. Transitional representations share case frame structuring and the use of variables with ILR, but they share sentence-type marking (declarative, interrogative, imperative) and indirect forms with SSR. Each transitional representation has an associated ILR, and a successful match of a SSR against the transitional representation (e.g., by the reference matching process) automatically binds the variables in the ILR.

Other Representations

It is quite possible that some semantic domains will require other sorts of representations. While the units of the current ILR are not to be construed as lexical items, ILR is intentionally linguistically oriented. There would no doubt be a need for quite different representations, e.g. tabular information and mathematical formulas. The representations summarized in this section, then, are seen as a minimum, not necessarily a complete set.

---

82. This representation type was called "subsurface level representation" in earlier versions of the model.

OFFICIAL DISTRIBUTION LIST

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314
12 copies

Office of Naval Research
Information Systems Program
Code 437
Arlington, VA 22217
2 copies

Office of Naval Research
Branch Office/Boston
Building 114, Section D
666 Summer Street
Boston, MA 02210
1 copy

Office of Naval Research
Branch Office/Chicago
536 South Clark Street
Chicago, IL 60605
1 copy

Office of Naval Research
Branch Office/Pasadena
1030 East Green Street
Pasadena, CA 91106
1 copy

New York Area
715 Broadway - 5th floor
New York, N. Y. 10003
1 copy

Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D. C. 20375
6 copies

Assistant Chief for Technology
Office of Naval Research
Code 200
Arlington, VA 22217
1 copy

Office of Naval Research
Code 455
Arlington, VA 22217
1 copy

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

Office of Naval Research
Code 458
Arlington, VA 22217
1 copy

Naval Ocean Systems Center,Code 91
Headquarters-Computer Sciences &
Simulation Department
San Diego, CA 92152
Mr. Lloyd Z. Maudlin
1 copy

Mr. E. H. Gleissner
Naval Ship Research & Development Center
Computation & Math Department
Bethesda, MD 20084
1 copy

Captain Grace M. Hopper (008)
Naval Data Automation Command
Washington Navy Yard
Building 166
Washington, D. C. 20374
1 copy

Mr. Kin B. Thompson
Technical Director
Information Systems Division
(OP-91T)
Office of Chief of Naval Operations
Washington, D. C. 20350
1 copy

Captain Richard L. Martin, USN
Commanding Officer
USS Francis Marion (LPA-249)
FPO New York, N. Y. 09501
1 copy