Two Constraints on Speech Act Ambiguity

Elizabeth A. Hinkelman and James F. Allen
Computer Science Department
The University of Rochester
Rochester, New York 14627

ABSTRACT

Existing plan-based theories of speech act interpretation do not account for the conventional aspect of speech acts. We use patterns of linguistic features (e.g. mood, verb form, sentence adverbials, thematic roles) to suggest a range of speech act interpretations for the utterance. These are filtered using plan-based conversational implicatures to eliminate inappropriate ones. Extended plan reasoning is available but not necessary for familiar forms. Taking speech act ambiguity seriously, with these two constraints, explains how "Can you pass the salt?" is a typical indirect request while "Are you able to pass the salt?" is not.

1. The Problem

Full natural language systems must recognize speakers' intentions in an utterance. They must know when the speaker is asserting, asking, or making a social or official gesture [Searle 69, Searle 75], in addition to its content. For instance, the ordinary sentence

(1) Can you open the door?

might in context be a question, a request, or even an offer. Several kinds of information complicate the recognition process. Literal meaning, lexical and syntactic choices, agents' beliefs, the immediate situation, and general knowledge about human behavior all clarify what the ordinary speaker is after.

Given an utterance and context, we model how the utterance changes the hearer's state. Previous work falls roughly into three approaches, each with characteristic weaknesses: the idiom approach, the plan based approach, and the descriptive approach.

The idiom approach is motivated by pat phrases like:

(2) a: Can you please X?
   b: Would you kindly X?
   c: I'd like X.
   d: May I X?
   e: How about X?

They are literally questions or statements, but often used as requests or in (e), suggestions. The system could look for these particular strings, and build the corresponding speech act using the complement as a parameter value. But such sentences are not true idioms, because the literal meaning is also possible in many contexts. Also, one can respond to the literal *and* nonliteral acts: "Sure, it's the 9th." The idiom approaches are too inflexible to choose the literal reading or to accommodate ambiguity. They lack a theory connecting the nonliteral and literal readings. Another problem is that some classic examples are not even pat phrases:

(3) a: It's cold in here.
   b: Do you have a watch on?

In context, (a) may be a request to close the window. Sentence (b) may be asking what time it is or requesting to borrow the watch. The idiom approach allows neither for context nor the reasoning connecting utterance and desired action.

The plan based approach [Allen 83, McCafferty 86, Perrault 80, Sidner 81] presumes a mechanism modelling human problem solving abilities, including reasoning about other agents and inferring their intentions. The system has a model of the current situation and the ability to choose a course of action. It can relate uttered propositions to the current situation: being cold in here is a bad state, and so you probably want me to do something about it; the obvious solution is for me to close the window, so, I understand, you mean for me to close the window. The plan based approach provides a tidy, independently motivated theory for speech act interpretation.

It does not use language-specific information, however. Consider

(4) a: Can you speak Spanish?
   b: Can you speak Spanish, please?

Here the addition of a word alters an utterance which is a yes/no question in typical circumstances to a request. This is not peculiar to "please":

(5) a: Can you open the door?
   b: Are you able to open the door?

Here two sentences, generally considered to have the same semantics, differ in force: the first may be a question, an offer, or a request, the second, only a question. Further, different languages realize speech acts in different ways: greetings, for example (or see [Horn 84]).

(6) a: You want to cook dinner.
   b: You wanna toss your coats in there?

The declarative sentence (a) can be a request, idiomatic to Hebrew, while the nearest American expression is interrogative (b). Neither is a request in British English. The plan based approach has nothing to say about these differences. Neither does it explain the psycholinguistic [Gibbs 84] finding that people access idiomatic interpretations in context more quickly than literal ones. Psycholinguistically plausible models cannot derive idiomatic meanings from literal meanings.

Descriptive approaches cover large amounts of data. [Brown 80] recognized the diversity of speech act phenomena and included the first computational model with wide coverage, but lacked theoretical claims and did not handle the language-specific cases well. [Gordon 75] expresses some very nice

generalizations, but lacks motivation and sufficient detail. It does not account for examples like numbers 3, 4, 6 or 7. In number 3, for example, one asks a question by asking literally whether the hearer knows the answer. A plan-based approach would argue that knowing the answer is a precondition for stating it, and this logical connection enables identification of the real question. But Gordon and Lakoff write off this one, because their sincerity conditions are inadequate.

We augment the plan-based approach with a linguistic component: compositional rules associating linguistic features with partial speech act descriptions. The rules express linguistic conventions that are often motivated by planning theory, but they also allow for an element of arbitrariness in just which forms are idiomatic to a language, and just which words and features mark it. For this reason, conventions of use cannot be handled directly by the plan reasoning mechanism. They require an interpretation process paralleling syntactic and semantic interpretation, with the same provisions for merging of partial interpretations and postponement of ambiguity resolution. The compositionality of partial speech act interpretations and use of ambiguity are both original to our approach.

Once the utterances have been interpreted by our conventional rules to produce a set of candidate conventional interpretations, these interpretations are filtered by the plan reasoner. Plan reasoning processes unconventional forms in the same spirit as earlier plan-based models, finding non-conventional interpretations and motivating many conventional ones. We propose a limited version of plan reasoning, based on an original claim about conversational implicature, which is adequate for filtering conventional interpretations.

Section 2 will explain the linguistic computation which interprets linguistic features as speech act descriptions. Section 3 describes plan reasoning techniques that are useful for speech act interpretation and presents our view of plan reasoning. Section 4 presents the overall process combining these two parts.

2. Linguistic Constraints

Speech act interpretation has many similarities to the plan recognition problem. Its goal is, given a situation and an utterance, to understand what the speaker was doing with that utterance, and to find a basis for an appropriate response. In our case this will mean identifying a set of plan structures representing speech acts, which are possible interpretations of the utterance. In this section we show how to use compositional, language-specific rules to provide evidence for a set of partial speech act interpretations, and how to merge them. Later, we use plan reasoning to constrain, supplement, and decide among this set.

2.1. Notational Aside

Our notation is based on that of [Allen 87]. Its essential form is (category <slot filler> <slot filler>...). Categories may be syntactic, semantic, or from the knowledge base. A filler may be a word, a feature, a knowledge-base object (referent) or another (category...) structure.

Two slots associated with syntactic categories may seem unusual: SEM and REF. They contain the unit's semantic interpretation, divided into two components. The SEM slot contains a structural-semantic representation of this instance, based on a small, finite set of thematic roles for verbs and noun phrases. It captures the linguistic generalities of verb subcategorization and noun phrase structure. Selectional restrictions, identification of referents, and other phenomena involving world knowledge are captured in the REF slot. It contains a translation of the SEM slot's logical form into a framelike knowledge representation language, in which richer and more specific role assignments can be made. SEM thematic roles correspond to different knowledge base roles according to the event class being described, and in REF the corresponding event and argument instances are identified if possible. Distinguishing logical form from knowledge representation is an experiment intended to clarify the notion of semantic roles in logical form, and to reduce the complexity of the interpretation process.

The sentence "Can you speak Spanish?" is shown below.

```
(S MOOD YES-NO-Q
   VOICE ACT
   SUBJ (NP HEAD you
             SEM (HUMAN h1)
             REF Suzanne)
   AUXS can
   MAIN-V speak
   TENSE PRES
   OBJ (NP HEAD Spanish
            SEM (LANG ID s1)
            REF ls1)

   SEM (CAPABLE
            TENSE PRES
            AGENT h1
            THEME (SPEAK OBJECT s1))

   REF (ABLE-STATE
            AGENT Suzanne
            ACTION (USE-LANGUAGE
                        AGENT Suzanne
                        LANG ls1)))
```

The outermost category is the syntactic category, sentence. It has many ordinary syntactic features, subject, object, and verbs. The subject is a noun phrase that describes a human and refers to a person

named Suzanne, the object a language, Spanish. The semantic structure concerns the capability of the person to speak a language. In the knowledge base, this becomes Suzanne's ability to use Spanish as a language.

2.2. Evidence for Interpretations

The utterance provides clues to the hearer, but we have already seen that its relation to its purpose may be complex. We need to make use of lexical and syntactic as well as semantic and referential information. In this section we will look at rules using all of these kinds of information, introducing the notation for rules as we go. Rules consist of a set of features on the left-hand side, and a set of partial speech act descriptions on the other. The rule should be interpreted as saying that any structure matching the left hand side *must* be interpreted as one of the speech acts indicated on the right hand side. The speech act descriptions themselves are also in (category <slot filler> ... <slot filler>) notation. Their categories are simply their types in the knowledge base's abstraction hierarchy, in which the category SPEECH-ACT abstracts all speech act types. Slot names and filler types also are defined by the abstraction hierarchy, but a given rule need not specify all slot values. Here is a lexical rule: the adverb "please" occurring in any syntactic unit signals a request, command, or other act in the directive class.

```
(? ADV please) =(1)=>
```

```
(DIRECTIVE-ACT)
```

Although this is a very simple rule, its correctness has been established by examination of some 43 million words of Associated Press news stories. This corpus contains several hundred occurrences of "please", the most common form being the preverbal adverb in a directive utterance.

A number of useful generalizations are based on the syntactic mood of sentences. As we use the term, mood is an aggregate of several syntactic features taking the values DECLARATIVE, IMPERATIVE, YES-NO-Q, WH-Q. Many different speech act types occur with each of these values, but in the absence of other evidence an imperative is likely to be a command and a declarative, an Inform. An interrogative sentence may be a question or possibly another speech act.

```
(S MOOD YES-NO-Q) =(2)=>
```

```
((ASK-ACT PROP V(REF))
 (SPEECH-ACT))
```

The value function V returns the value of the specified slot of the sentence. Thus rule 2 has the proposition slot PROP filled with the value of the REF slot of the sentence. It matches sentences whose

mood is that of a yes/no question, and interprets them as asking for the truth value of their explicit propositional content. Thus matching this rule against the structure for "Can you speak Spanish?" would produce the interpretations

```
((ASK-ACT
    PROP (ABLE-STATE AGENT Suzanne
                ACTION (USE-LANGUAGE
                         AGENT Suzanne
                         LANG ls1)))
```

```
 (SPEECH-ACT))
```

Interrogative sentences with modal verbs and a subject "you" are typically requests, but may be some other act:

```
(S MOOD YES-NO-Q    =(3)=>
    VOICE ACT
    SUBJ (NP PRO you)
    AUXS {can could will would might}
    MAIN-V +action)
```

```
((REQUEST-ACT ACTION V(ACTION REF))
 (SPEECH-ACT))
```

Rule 3 interprets "Can you...?" questions as requests, looking for the subject "you" and any of these modal verbs. Lists in curly brackets (e.g. {can could will would might}) signify disjunctions; one of the members must be matched. In this rule, the value function V follows a chain of slots to find a value. Thus V(ACTION REF) is the value of the ACTION slot in the structure that is the value of the REF slot. Note that an unspecified speech act is also included as a possibility in both rules. This is because it is also possible that the utterance might have a different interpretation, not suggested by the mood.

Some rules are based in the semantic level. For example, the presence of a benefactive case may mark a request, or it may simply occur in a statement or question.

```
(S MAIN-V +action
    SEM (? BENEF ?)) =(4)=>
```

```
((DIRECTIVE-ACT ACT V(REF))
 (SPEECH-ACT))
```

Recall that we distinguish the semantic level from the reference level, inasmuch as the semantic level is simplified by a strong theory of thematic roles, or cases, a small standard set of which may prove adequate to explain verb subcategorization phenomena [Jackendoff 72] The reference level, by

contrast, is the language of the knowledge base, in which very specific domain roles are possible. To the extent that referents can be identified in the knowledge base (often as skolem functions) they appear at the reference level. This rule says that any way of stating a desire may be a request for the object of the want.

```
(S MOOD DECL =(5)=>
   VOICE ACT
   TENSE PRES
   REF (WANT-ACT ACTOR !s))

(REQUEST-ACT
   ACT V(DESID WANT-ACT REF))
```

It will match any sentence that can be interpreted as asserting a want or desire of the agent, such as

    (7) a: I need a napkin.
        b: I would like two ice creams.

The object of the request is the WANT-ACT's desideratum. (The desideratum is already filled by reference processing.) One may prefer an account that handles generalizations from the REF level by plan reasoning; we will discuss this point later. For now, it is sufficient to note that rules of this type are capable of representing the conventions of language use that we are after.

2.3. Applying the Rules

We now consider in detail how to apply the rules. For now, assume that the utterance is completely parsed and semantically interpreted, unambiguously, like the sentence "Can you speak Spanish?" as it appeared in Sect. 2.1.

Interpretation of this sentence begins by finding rules that match with it. The matching algorithm is a standard unification or graph matcher. It requires that the category in the rule match the syntactic structure given. All slots present in the rule must be found on the category, and have equal values, and so on recursively. Slots not present in the rule are ignored. If the rule matches, the structures on the right hand side are filled out and become partial interpretations. We need a few general rules to fill in information about the conversation:

```
(?)   =(6)=>   ((SPEECH-ACT AGENT !s))
```

Rule 6 says that an utterance of any syntactic category maps to a speech act with agent specified by the global variable !s. (The processes of identifying speaker and hearer are assumed to be contextually defined.) The partial interpretation it yields for the Spanish sentence is a speech act with agent Mrs. de Prado:

```
((SPEECH-ACT AGENT Mrs. de Prado))
```

The second rule is analogous, filling in the hearer.

```
(?)   =(7)=>   ((SPEECH-ACT HEARER !h))
```

For our example sentence, it yields a speech act with hearer Suzanne.

```
((SPEECH-ACT HEARER Suzanne))
```

Rule no. 2 given earlier, for yes/no questions, produces these interpretations:

```
((ASK-ACT
   PROP (ABLE-STATE
         AGENT Suzanne
         ACTION (USE-LANGUAGE
                 AGENT Suzanne
                 LANG ls1)))

  (SPEECH-ACT))
```

The indirect request comes from rule no. 3 above. To apply it, we match the subject "you" and the modal auxiliary "can", and the features of yes/no mood and active voice.

```
((REQUEST-ACT
   ACTION (USE-LANGUAGE
           AGENT Suzanne
           LANG ls1)))

  (SPEECH-ACT))
```

We now have four sets of partial descriptions, which must be merged.

2.4. Combining Partial Descriptions

The combining operation can be thought of as taking the cross product of the sets, merging partial interpretations within each resulting set, and returning those combinations that are consistent internally. We expect that psycholinguistic studies will provide additional constraints on this set, e.g. commitment to interpretations triggered early in the sentence.

The operation of merging partial interpretations is again unification or graph matching; when the operation succeeds the result contains all the information from the contributing partial interpretations. The cross product of our first two sets is simple; it is the pair consisting of the interpretation for speaker and hearer. These two can be merged to form a set containing the single speech act with speaker Mrs. de Prado and hearer Suzanne. The cross product of this with the results of the mood rule contains two pairs. Within the first pair, the ASK-ACT is a subtype of SPEECH-ACT and

therefore matches, resulting in a request with the proper speaker and hearer. The second pair results in no new information, just the SPEECH-ACT with speaker and hearer. (Recall that the mood rule must allow for other interpretations of yes/no questions, and here we simply propagate that fact.)

Now we must take the cross product of two sets of two interpretations, yielding four pairs. One pair is inconsistent because REQUEST-ACT and ASK-ACT do not unify. The REQUEST-ACT gets speaker and hearer by merging with the SPEECH-ACT, and the ASK-ACT slides through by merging with the other SPEECH-ACT. Likewise the two SPEECH-ACTs match, so in the end we have an ASK-ACT, REQUEST-ACT, and the simple SPEECH-ACT.

```
((REQUEST-ACT
     AGENT Mrs. de Prado
     HEARER Suzanne
     ACTION (USE-LANGUAGE
                 AGENT Suzanne
                 LANG ls1)))

(ASK-ACT
  AGENT Mrs. de Prado
  HEARER Suzanne
  PROP (ABLE-STATE
             AGENT Suzanne
             ACTION (USE
                     AGENT Suzanne
                     OBJECT ls1)))

(SPEECH-ACT AGENT Mrs. de Prado)
             HEARER Suzanne))
```

At this stage, the utterance is ambiguous among these three interpretations. Consider their classifications in the speech act hierarchy. The third abstracts the other two, and signals that there may be other possibilities, those it also abstracts. Its significance is that it allows the plan reasoner to suggest these further interpretations, and it will be discussed later. If there are any expectations generated by top-down plan recognition mechanisms, say, the answer in a question/answer pair, they can be merged in here.

2.5. Further Linguistic Considerations

We have used a set of compositional rules to build up multiple interpretations of an utterance, based on linguistic features. They can incorporate lexical, syntactic, semantic and referential distinctions. Why does the yes/no question interpretation seem to be favored in the Spanish example? We hypothesize that for utterances taken out of context, people make pure frequency judgements. And questions about one's language ability are much more common than requests to speak one. Such a single-utterance request is possible only in contexts where the intended *content* of the Spanish-speaking is clear or

clearly irrelevant, since "speak" doesn't subcategorize for this crucial information. (cf. "Can you read Spanish? I have this great article ....")

The statistical base can be overridden by lexical information. Recall 5(b) "Can you speak Spanish, please?" The "please" rule (above) yields only the request interpretation, and fails to merge with the ASK-ACT. It also merges with the SPEECH-ACT, but the result is again a request, merely adding the possibility that the request could be for some other action. No such action is likely to be identified. The "please" rule is very strong, because it can override our expectations. The final interpretations for "Can you speak Spanish, please?" do not include the literal interpretation:

```
((REQUEST-ACT
     AGENT Mrs. de Prado
     HEARER Suzanne
     ACTION (USE-LANGUAGE
                 AGENT Suzanne
                 LANG ls1)))

((REQUEST-ACT AGENT Mrs. de Prado
             HEARER Suzanne)
```

Here Suzanne is probably being asked to continue the present dialogue in Spanish.

Some linguistic features are as powerful as "please", as can be seen by the incoherence of the following, where each sentence contains contradictory features.

(8) a: *Should you go home, please?
    b: *Shouldn't you go home, please?
    c: *Why not go home, please?

Modal verbs can be quite strong, and intonation as well. Other features are more suggestive than definitive. The presence of a benefactive case (rule above) may be evidence for an offer or request, or just happen to appear in an inform or question. Sentence mood is weak evidence: it is often overridden, but in the absence of other evidence it it becomes important. The adverbs "kindly" and "possibly" are also weak evidence for a request, and large class of sentential adverbs is associated primarily with Inform acts.

(9) a: *Unfortunately, I promise to obey orders.
    b: Surprisingly, I'm leaving next week.
    c: Actually, I'm pleased to see you.

Explicit performative utterances [Austin 62] are declarative, active, utterances whose main verb identifies the action explicitly. The sentence meaning corresponds exactly to the action performed.

```
(S MOOD DECL
   VOICE ACT
   SUBJ (NP HEAD i)
   MAIN-V +performative
   TENSE PRES)
```

= (8) =>   V(REF)

Note that the rule is not merely triggering off a keyword. Presence of a performative verb *without* the accompanying syntactic features will not satisfy the performative rule.

2.6. The Limits of Conventionality

We do not claim that all speech acts are conventional. There are variations in convention across languages, of course, and dialects, but idiolects also vary greatly. Some people, even very cooperative ones, do not recognize many types of indirect requests. Too, there is a form of request for which the generalization is obvious but only special cases seem idiomatic:

(10) a: Got a light?
     b: Got a dime?
     c: Got a donut? (odd request)
     d: Do you have the time?
     e: Do you have a watch on?

There are other forms for which the generalization is obvious but no instance seems idiomatic: if someone was assigned a task, asking whether it's done is as good as a request.

(11) Did you wash the dishes?

In the next examples, there is a clear logical connection between the utterance and the requested action. We could write a rule for the surface pattern, but the rule is useless because it cannot verify the logical connection. This must be done by plan reasoning, because it depends on world knowledge. The first sentences can request the actions they are preconditions of. The second set can request actions they are effects of. Because these requests operate via the conditions on the domain plan rather than the speech act itself, they are beyond the reach of theories like Gordon&Lakoff 's, which have very simple notions of what a sincerity condition can be.

(12) a: Is the garage open?
     b: Did the dryer stop?
     c: The mailman came.
     d: Are you planning to take out the garbage?

(13) a: Is the car fixed?
     b: Have you fixed the car?
     c: Did you fix the car?

Plan reasoning provides an account for all of these examples. The fact that certain examples can be handled by either mechanism we regard as a strength of the theory: it leads to robust natural language processing systems, and explains why "Can you X?" is such a successful construction. Both mechanisms work well for such utterances, so the hearer has two ways to understand it correctly. These last examples, along with "It's cold in here", really require plan reasoning.

3. Role of Plan Reasoning

Plan reasoning constitutes our second constraint on speech act recognition. There are four roles for plan reasoning in the recognition process. Specifically, plan reasoning

1) eliminates speech act interpretations proposed by the linguistic mechanism, if they contradict known intentions and beliefs of the agent.

2) elaborates and makes inferences based on the remaining interpretations, allowing for non-conventional speech act interpretations.

3) can propose interpretations of its own, when there is enough context information to guess what the speaker might do next.

4) provides a competence theory motivating many of the conventions we have described.

Plan reasoning rules are based on the causal and structural links used in plan construction. For instance, in planning start with a desired goal proposition, plan an action with that effect, and then plan for its preconditions. There are also recognition schemas for attributing plans: having observed that an agent wants an effect, believe that they may plan an action with that effect, and so on. For modelling communication, it is necessary to complicate these rules by embedding the antecedent and consequent in one-sided mutual belief operators [Allen 83]. In the Allen approach, our Spanish example hinges on the acts' preconditions: Suzanne will not attribute a question to Mrs. de Prado if she believes she already knows the answer, but this knowledge could be the basis for a request. Sentences like "It's cold in here" are also interpreted by extended reasoning about the intentions an agent could plausibly have. We use extended reasoning for difficult cases, and the more restricted plan-based conversational implicature heuristic [Hinkelman 87], [Hinkelman forthcoming] as a plausibility filter adequate for most common cases.

4. Two Constraints Integrated

Section 2 showed how to compute a set of possible speech act interpretations compositionally, from conventions of language use. Section 3 showed how plan reasoning, which motivates the conventions, can be used to further develop and restrict the interpretations. The time has come to integrate the two into a complete system.

4.1. Interaction of the Constraints

The plan reasoning phase constrains the results of the linguistic computation by eliminating interpretations, and reinterpreting others. The linguistic computation constrains plan reasoning by providing the input; the final interpretation must be in the range specified, and only if there is no plausible interpretation is extended inference explicitly invoked. Recall that the

217

linguistic rules control ambiguity: because the right hand side of the rule must express *all* the possibilities for this pattern, a single rule can limit the range of interpretations sharply. Consider

(14) a: I hereby inform you that it's cold in here.
    b: It's cold in here.

The explicit performative rules, triggered by "hereby" and by a performative verb in the appropriate syntactic context, each allow for only an explicit performative interpretation. (a) is unambiguous, and if it is consistent with context no extended reasoning is needed for speech act identification purposes. (In fact the hearer will probably find the formality implausible, and try to explain that.) By contrast, the declarative rule proposes two speech acts for (b), the Inform and the generic speech act. The ambiguity allows the plan reasoner to identify other interpretations, particularly if in context the Inform interpretation is implausible.

The entire speech act interpretation process is now as follows. Along with the usual compositional linguistic processes, we build up and merge hypotheses about speech act interpretations. The resulting interpretations are passed to the implicature module. The conversational implicatures are computed, discounting interpretations if they are in conflict with contextual knowledge. If a plausible, non-contradictory interpretation results, it can be accepted. Allen-style plan reasoning is invoked to identify the speech act only if remaining ambiguity interferes with planning or if no completely plausible interpretations remain. After that, plan reasoning may proceed to elaborate the interpretation or to plan a response.

Consider the central example of this paper. Three interpretations were were proposed for "Can you speak Spanish?", in Section 2.

As they become available, the next step in processing is to check plausibility by attempting to verify the act's conversational implicatures. We showed how the Ask act is ruled out by its implicatures, when the answer is known. Likewise, in circumstances where Suzanne is known not to speak Spanish, the Request is eliminated.

The generic speech act is present under most circumstances, but adds little information except to allow for other possibilities. Because in any of these contexts a specific interpretation is acceptable, no further inference is necessary for identifying the speech act. If it is merely somewhat likely that Suzanne speaks Spanish, both specific interpretations are possible and both may even be intended by Mrs. de Prado. Further plan reasoning may elaborate or eliminate possibilities, or plan a response. But it is not required for the main effort of speech act identification.

4.2. The Role of Ambiguity

If no interpretations remain after the plausibility check, then the extended plan reasoning may be invoked to resolve a possible misunderstanding or mistaken belief. If several remain, it may not be necessary to disambiguate. Genuine ambiguity of intentions is quite common in speech and often not a problem. For instance, the speaker may mention plans to go to the store, and leave unclear whether this constitutes a promise.

In cases of genuine ambiguity, it is possible for the hearer to respond to each of the proposed interpretations, and indeed, politeness may even require it. Consider (b)-(g) as responses to (a).

(15) a: Do you have our grades yet?
    b: No, not yet.
    c: Yes, I'm going to announce them in class.
    d: Sure, here's your paper. (hands paper.)
    e: Here you go. (hands paper.)
    f: *No.
    g: *Yes.

The main thing to note is that it is infelicitous to ignore the Request interpretation; the polite responses acknowledge that the speaker wants the grades.

Note that within the framework of "speaker-based" meaning, we emphasize the role of the hearer in the final understanding of an utterance. An important point is that while the speech act *attempted* depends on the speaker's intentions, the speech act *accomplished* also depends on the hearer's ability to recognize the intentions, and to some extent their own desires in the matter. Consider an example from [Clark 88]:

(16) a: Have some asparagus.
    b: No, thanks.

(17) a: Have some asparagus.
    b: OK, if I have to....

The first hearer treats the sentence as an offer, the second as a command. If the speaker intended otherwise, it must be corrected quickly or be lost.

4.3. The Implementation

Our system is implemented using common lisp and the Rhetorical knowledge representation system [Miller 87], which provides among other things a hierarchy of belief spaces. The linguistic speech act interpretation module been implemented, with merging, as well as the implicature calculation and checking module. So given the appropriate contexts, the Spanish example runs. Extended plan reasoning will eventually be added.

There are of course open problems. One would like to experiment with large interpretation rule sets, and with the constraints from other modules. The projection problem, both for conversational implicature and for speech act interpretation, has not been examined directly. If a property like conversational implicature or presupposition is

computed at the clause level, one wants to know whether the property survives negation, conjunction, or any other syntactic embedding. [Horton 87] has a result for projection of presuppositions, which may be generalizable. The other relevant work is [Hirschberg 85] and [Gazdar 79]. Plan recognition for discourse, and the processing of cue words, are related areas.

5. Conclusion

To determine what an agent is doing by making an utterance, we must make use of not only general reasoning about actions in context, but also the linguistic features which by convention are associated with specific speech act types. To do this, we match patterns of linguistic features as part of the standard linguistic processing. The resulting partial interpretations are merged, and then filtered by determining the plausibility of their conversational implicatures. Assuming no errors on the part of the speaker, the final interpretation is constrained to lie within the range so specified.

If there is not a plausible interpretation, full plan reasoning is called to determine the speaker's intentions. Remaining ambiguity is not a problem but simply a more complex basis for the hearer's planning processes. Linguistic patterns and plan reasoning together constrain speech act interpretation sufficiently for discourse purposes.

Acknowledgements

References

[Allen 83] Allen, J., "Recognizing Intentions From Natural Language Utterances," in *Computational Models of Discourse*, Brady, M. and Berwick, B. (ed.), MIT Press, Cambridge, MA, 1983, 107-166.

[Allen 87] Allen, J., *Natural Language Understanding*, Benjamin Cummings Publishing Co., 1987.

[Austin 62] Austin, J. L., *How to Do Things with Words*, Harvard University Press, Cambridge, MA, 1962.

[Brown 80] Brown, G. P., "Characterizing Indirect Speech Acts," *American Journal of Computational Linguistics* 6:3-4, July-December 1980, 150-166.

[Clark 88] Clark, H., Collective Actions in Language Use, Invited Talk, September 21, 1988.

[Gazdar 79] Gazdar, G., *Pragmatics: Implicature, Presupposition and Logical Form*, Academic Press, New York, 1979.

[Gibbs 84] Gibbs, R., "Literal Meaning and Psychological Theory," *Cognitive Science 8*, 1984, 275-304.

[Gordon 75] Gordon, D. and Lakoff, G., "Conversational Postulates," in *Syntax and Semantics V. 3*, Cole, P. and Morgan, J. L. (ed.), Academic Press, New York, 1975.

[Hinkelman 87] Hinkelman, E., "Thesis Proposal: A Plan-Based Approach to Conversational Implicature," TR 202, Dept. Computer Science, University of Rochester, June 1987.

[Hirschberg 85] Hirschberg, J., "A Theory of Scalar of Implicature," MS-CIS-85-56, PhD Thesis, Department of Computer and Information Science, University of Pennsylvania, December 1985.

[Horn 84] Horn, L. R. and Bayer, S., Short-Circuited Implicature: A Negative Contribution, Vol. 7, 1984.

[Horton 87] Horton, D. L., "Incorporating Agents' Beliefs in a Model of Presupposition," Technical Report CSRI-201, Computer Systems Research Institute, University of Toronto, Toronto, Canada, August 1987.

[Jackendoff 72] Jackendoff, R. S., *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, 1972.

[McCafferty 86] McCafferty, A. S., Explaining Implicatures, 23 October 1986.

[Miller 87] Miller, B. and Allen, J., *The Rhetorical Knowledge Representation System: A User's Manual*, forthcoming technical report, Department of Computer Science, University of Rochester, 1987.

[Perrault 80] Perrault, C. R. and Allen, J. F., "A Plan-Based Analysis of Indirect Speech Acts," *American Journal of Computational Linguistics* 6:3-4, July-December 1980, 167-82.

[Searle 69] Searle, J., in *Speech Acts*, Cambridge University Press, New York, 1969.

[Searle 75] Searle, J., "Indirect Speech Acts," in *Syntax and Semantics, v3: Speech Acts*, Cole and Morgan (ed.), Academic Press, New York, NY, 1975.

[Sidner 81] Sidner, C. L. and Israel, D. J., "Recognizing Intended Meaning and Speakers' Plans," *Proc. IJCAI '81*, 1981, 203-208.

219