USING LINGUISTIC, WORLD, AND CONTEXTUAL KNOWLEDGE IN A PLAN RECOGNITION MODEL OF DIALOGUE[1]

LYNN LAMBERT and SANDRA CARBERRY
Department of Computer and Information Sciences
University of Delaware
Newark, Delaware 19716, USA
email: lambert@cis.udel.edu, carberry@cis.udel.edu

Abstract

This paper presents a plan-based model of dialogue that combines world, linguistic, and contextual knowledge in order to recognize complex communicative actions such as expressing doubt. Linguistic knowledge suggests certain discourse acts, a speaker's beliefs, and the strength of those beliefs; contextual knowledge suggests the most coherent continuation of the dialogue; and world knowledge provides evidence that the applicability conditions hold for those discourse acts that capture the relationship of the current utterance to the discourse as a whole.

1   Introduction

Recognizing the roles that utterances play in a dialogue and how the utterances should be interpreted in the context of preceding dialogue is a crucial part of a robust model of understanding. In order to perform this recognition, our tripartite plan-based model of dialogue identifies not only domain and problem-solving actions but also discourse or communicative actions that determine how utterances relate to each other. For this communicative action recognition, we combine information gleaned from a variety of knowledge sources: contextual, linguistic, and world knowledge. The combination of these different knowledge sources enables the recognition of complex communicative actions such as expressing doubt. Although our tripartite model recognizes three different kinds of actions (domain, problem-solving, and discourse), the focus of this paper will be the recognition of discourse actions and how a combination of knowledge sources enables us to perform this recognition.

2   Our Tripartite Model

A number of researchers have contended that a coherent discourse consists of segments that are related to one another through some type of structuring relation [Gri75, MT83] or have modeled discourse based on the semantic relationship of individual clauses [Pol86] or groups of clauses [Rei78]. But all of these fail to capture the goal-oriented nature of discourse. Grosz and Sidner [GS86] argue that recognizing the structural relationships among the intentions underlying a

discourse is necessary to identify discourse structure; although they do not provide the details of a computational mechanism for recognizing these relationships, they do argue convincingly that it requires multiple knowledge sources. We have developed a plan-based model of dialogue and have incorporated into our model Grosz and Sidner's claim that linguistic, contextual, and world knowledge should be combined in recognizing the role of an utterance in a discourse.

We contend that at least three kinds of actions, domain, problem-solving, and discourse, should be captured by a model of task-oriented dialogue. Many researchers [All79, Car87, LA87, Sid85, GS86] have demonstrated that recognition of domain actions endows a system with the ability to successfully address many important and difficult problems in understanding. Several researchers have also investigated the recognition of problem-solving actions [LA87, Ram91, Wil81]. For example, if a user wants to earn a degree, the user might perform problem-solving actions of 1) evaluating alternative degrees (i.e., the user might decide whether a BS or a BA is more desirable), 2) instantiating the type of degree to be earned, and 3) building a plan for performing the domain action of earning the selected degree.

Carberry [Car89] points out the importance of recognizing discourse actions, the communicative actions that speakers perform in making an utterance (e.g., asking a question, providing background information, or expressing surprise). Discourse actions provide expectations for subsequent utterances (e.g., when a question is asked, one expects the question to be accepted and eventually answered). Recognition of some discourse actions such as *Give-Background* also explains the purpose of an utterance and how it should be interpreted; rather than just a statement of fact, the utterance providing background information should be used by the system to fill in necessary background knowledge in order to fully understand related utterances.

We capture these three types of actions in separate levels of our discourse model, which we refer to as a DM [LC91]. Within each of these levels, actions may contribute to other actions on the same level; for example, on the discourse level, providing background data, asking a question, and answering a question all can be part of obtaining information. Thus, actions at each level form a tree structure in which each node represents an action that a participant is performing and the children of a node represent actions pursued in order to perform the parent action. However, discourse, problem-solving, and domain actions are not
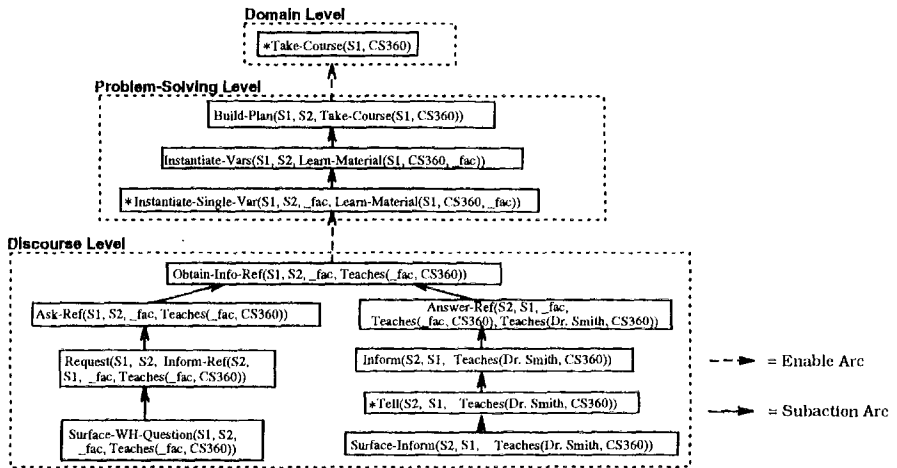
Figure 1: Dialogue Model for two utterances

completely independent of one another; discourse actions may be executed in order to obtain the information necessary for performing a problem-solving action and problem-solving actions may be executed in order to construct a domain plan. Our model captures this interaction by allowing links between the actions at adjacent levels. Figure 1 contains a sample DM derived from two utterances, and section 3 describes how the DM in Figure 1 is constructed.

3   Different Kinds of Knowledge

The following dialogue will be used to demonstrate why our system needs world, contextual, and linguistic knowledge, and to show how the combination of these different knowledge sources enables the system to recognize implicit acceptance of previously communicated propositions and to identify the role of utterances that cannot be determined from one or two knowledge sources alone. The system is playing the role of S2.

(1)   S1:   *Who is teaching CS360?*
(2)   S2:   *Dr. Smith is teaching CS360.*
(3)   S1:   *But isn't CS360 an undergrad course?*
(4)   S2:   *Yes.*
(5)   *Dr. Smith teaches graduate*
             and undergrad courses.
(6)   S1:   *Who handles the CS360 lab?*

3.1   World Knowledge

In order to recognize how actions are intended to contribute to other actions, our system needs knowledge about how to perform actions. This world knowledge is provided in the form of a library of discourse, problem-solving, and domain recipes [Pol90]. Our representation of a recipe includes a header giving the name of the recipe and the action that it accomplishes, preconditions, applicability conditions, constraints, a body, effects, and a goal. Applicability conditions represent conditions that must be satisfied in order for the recipe to be reasonable to apply in the given situation whereas constraints limit the allowable instantiation of variables in each of the components of a recipe [LA87, Car87]. Figure 2 contains a sample discourse recipe.

Given the semantic representation of a new utterance, the system must assimilate the utterance and produce an updated dialogue model (DM). Plan inference rules [All79] and constraint satisfaction [LA87, Car87] suggest chains of higher level actions that an utterance may be part of, and focusing heuristics [Car87, Sid81] order these inference paths according to coherence. For example, the semantic representation of (1) is:

Surface-WH-Question(S1, S2, _fac, Teaches(_fac, CS360))

From this surface question, plan inference rules suggest that (1) is executing a *Request* action and that this *Request* action is part of an *Ask-Ref* action which in turn is part of an *Obtain-Info-Ref* action since each of these actions is part of the body of a recipe that performs the higher level action. As the system infers these actions, the system also tentatively ascribes certain beliefs that must hold in order for the agent to be pursuing these discourse actions. For example, in order for (1) to be part of an *Obtain-Info-Ref* action, S1 must not know the answer to the question; if S1 knew who was teaching CS360, this utterance might be part of a *Test-Listener* action instead. These requisite beliefs are captured in the applicability conditions of discourse recipes. As the system infers actions, it must be plausible that the applicability conditions are

Discourse Recipe-C1: {_agent1 expresses doubt to _agent2 about _prop1 because _agent1 believes _prop2 to be true}

Action:	Express-Doubt(_agent1, _agent2, _prop1, _prop2, _rule)
App Cond:	believe(_agent1, _prop2)
	believe(_agent1, believe(_agent2, _prop1))
	believe(_agent1, _rule)
	believe(_agent1, ((_prop2 ∧ _rule) ⇒ ¬_prop1))
	in-focus(_prop1)
Body:	Convey-Uncertain-Belief(_agent1, _agent2, _prop2)
	Address-Q-Acceptance(_agent2, _agent1, _prop2)
Effects:	believe(_agent2, ¬believe(_agent1, _prop1))
	believe(_agent2, want(_agent1, Resolve-Conflict(_agent2, _agent1, _prop1, _prop2)))
Goal:	want(_agent2, Resolve-Conflict(_agent2, _agent1, _prop1, _prop2))

Figure 2: Sample Discourse Recipe

satisfied; otherwise, the inference is rejected.

So, another part of our system's world knowledge is a model of the speaker's beliefs. Since our investigation of naturally occurring dialogues indicates that people express shades of belief in propositions and expect others to recognize these beliefs, we maintain a multi-strength model of beliefs to represent an agent's varying degrees of belief in a proposition, ranging from having no idea whether a proposition is true (or false), to being certain that a proposition is true (or false). We also maintain a model of a stereotypical user whose beliefs may be attributed to the speaker as appropriate during the course of the conversation.

After the system has inferred actions on the discourse level, it must identify how these relate to problem-solving and domain actions. This is accomplished by chaining between actions on adjacent levels of the DM. For example, once the system infers that (1) contributes to an *Obtain-Info-Ref* action on the discourse level, plan inference rules suggest that S1 wants the goal of the *Obtain-Info-Ref* action, namely knowref(S1, _fac, teaches(_fac, CS360)). Since knowing possible fillers for a parameter is a precondition to instantiating that parameter, the system infers that S1 wants to know who is teaching CS360 in order to instantiate the instructor parameter in an action to learn the material for that course; that is, the system infers that S1 wants *Instantiate-Single-Var(S1, S2, _fac, Learn-Material(S1, CS360, _fac))*. Since instantiating one parameter in an action is part of a plan to instantiate all of the parameters in that action, the system infers that S1 wants *Instantiate-Vars(S1, S2, Learn-Material(S1, CS360, _fac))*, and since this latter action is part of a recipe for building a plan, the system then infers the problem-solving action, *Build-Plan(S1, S2, Take-Course(S1, CS360))*. These Instantiate-Single-Var, Instantiate-Vars, and Build-Plan actions are entered into the problem-solving level of the DM. Building a plan to perform some domain action is a precondition to doing that action (assuming agents are acting intentionally), so the system infers that S1 wants *Take-Course(S1, CS360)*, and this domain action is entered into the domain level of the DM.

Once the system has assimilated an utterance into the DM, it must update its belief model for the speaker to reflect the beliefs that were tentatively ascribed to the speaker during the plan inference process. These beliefs can then be used in understanding subsequent utterances.

3.2 Contextual Knowledge

Each new utterance must be interpreted with respect to the existing dialogue context [GS86, Car87]. This process requires contextual knowledge, which our system captures with the use of the DM, a focus of attention which designates the most salient action on each level of the DM, and focusing heuristics which suggest the most coherent continuation of the dialogue. For example, on the discourse level, utterances that contribute to the currently focused action are more expected, and thus more coherent, than utterances that contribute to an ancestor that is further removed from the focus of attention. This contextual knowledge creates expectations that help determine how to interpret new utterances. For example, after a question is asked, the context suggests that acceptance of the question will be pursued (i.e., the listener will ensure that the question is understood, justified, and answerable); then it is expected that the question will be answered.

Because our system does not yet include a planner, we incorporate S2's utterances into the DM in the same way as S1's have been. So, continuing with our example dialogue, we express the semantic representation of (2) as:

Surface-Inform(S2, S1, Teaches(Dr. Smith, CS360))

Plan inference rules suggest that the surface inform might be part of a *Tell* action which might be part of an *Inform* action[2] which might be part of an *Answer-Ref* action which might in turn be part of an *Obtain-Info-Ref* action since each of these actions is part of the body of a recipe that performs the higher level action. Contextual knowledge is then used to determine how to relate (2) to the previous dialogue. Focusing heuristics suggest that the best interpretation of (2) is that it is part of a plan for performing the *Obtain-Info-Ref* action that was an ancestor of the *Request* action of utterance (1). No new problem-solving or domain actions are inferred.

Figure 1 gives the DM that our system builds from utterances (1) and (2) with the current focus of

---

[2] We differentiate between *telling* a listener some string of words and *informing* a listener of a proposition. In order to inform a listener of some proposition, the listener must first understand the content of the proposition; this is the goal of the *Tell* action. The goal of the *Inform* action is that the listener believe the communicated proposition.

attention on each level marked with an asterisk. Thus, we have seen that both world knowledge (consisting of a plan library and beliefs about the speaker's beliefs) and contextual knowledge (consisting of the existing DM, the current focus of attention, and focusing heuristics) are required in order to determine what actions a speaker is performing and how these actions relate to the previous dialogue.

3.3 Linguistic Knowledge

Linguistic knowledge must also be taken into account in recognizing the actions that a speaker is performing. This knowledge includes the surface form of an utterance and clue words. The surface form of an utterance is one way that a speaker communicates varying degrees of belief in a proposition. Consider, for example, the following two utterances:

    (7) *Is Dr. Smith teaching CS310?*
    (8) *Isn't Dr. Smith teaching CS310?*

The form of the utterance in (8) indicates an uncertain belief in the proposition that Dr. Smith is teaching CS310; utterance (7), however, conveys only a lack of knowledge about the proposition. Similarly (3) is not merely a Yes/No question; instead, this surface form conveys that S1 *thinks* that CS360 is an undergrad course, but is not certain of it. Our system uses the form of the utterance to recognize the strength of a speaker's beliefs; these beliefs are then used to determine whether the applicability conditions for the suggested discourse actions are satisfied.

The second type of linguistic information that we use is clue words. These linguistic clues often suggest what type of discourse action the speaker might be pursuing[LA87, Hin89].[3] We use these linguistic clues as evidence for discourse actions. For example, utterance (3) contains the clue word "but," which suggests a non-acceptance discourse action. Thus, the linguistic information that our system captures includes knowledge about the surface form of an utterance and about clue words.[4]

4 Combining Knowledge for Discourse Act Recognition

Because there are many ways that an utterance can continue a dialogue and because the correct interpretation is not always the one most strongly suggested by plan chaining and focusing heuristics, evidence from other knowledge sources is needed to identify the intended relationship between an utterance and the existing dialogue context. For example, the interpretation of (3) most strongly suggested by focusing heuristics is that of requesting clarification of (2) in order to understand it. Intuitively, however, (3) seems to be expressing doubt at S2's answer, not trying to understand

---

[3]The surface form of some utterances may also serve this purpose.
[4]The utterance itself in fact contains more information than just clue words, surface form, and propositional content, but our system uses only these three. Part of our future work includes incorporating other linguistic information such as tense.

it. Plan inference rules and focusing heuristics then are not sufficient to determine what role (3) is serving in the discourse. More knowledge is needed from other sources.

Intuitively, for each new utterance that relates to the previous dialogue on the discourse level,[5] there is some discourse action that captures this relationship and serves to describe the role of the new utterance with respect to the preceding dialogue. Since many such relationships are plausible (e.g., (3) could be interpreted as expressing doubt or as indicating a lack of understanding), we contend that evidence is required for recognizing the discourse action that identifies the correct relationship.

In our model, this discourse action will be an element of an inference path from the utterance to some action in the current tree structure on the discourse level. Furthermore, it will introduce new parameters which must be instantiated based on values from the DM in order for the path to terminate with an action already in the DM. By replacing these new parameters with values from the DM that are not present in the semantic representation of the utterance, we are hypothesizing a relationship between the new utterance and the existing discourse level of the DM. Thus, this action serves the aforementioned role of capturing how the new utterance relates to the current dialogue context. We will refer to such actions as *e-actions* since we contend that there must be evidence to support the inference of these actions.

For example, suppose that the semantic representation of an utterance such as (8) is

    Surface-Neg-YN-Question(S1, S2, PropA)

and that chaining suggests that the utterance is part of a *Convey-Uncertain-Belief* action which is part of a recipe for the *Express-Doubt* action shown in Figure 2. The parameters _agent1, _agent2, and _prop2 in the *Express-Doubt* action will be instantiated with the values S1, S2, and PropA that appear in the semantic representation of the utterance and propagate during chaining to the *Convey Uncertain-Belief* action and in turn to the *Express-Doubt* action. However, the parameters _prop1 and _rule are introduced for the first time in the *Express-Doubt* action and have many plausible instantiations. Continued chaining from the *Express-Doubt* discourse action could eventually lead to an *Inform* action, and we might equate this *Inform* with an *Inform* that already exists in the DM, thereby interpreting the new utterance as related to this previously identified action. Unifying the *Inform* action on the inference path with an *Inform* action in the DM, and propagating the resultant substitutions back down the inference chain, will result in _prop1 and _rule being instantiated based on information from the DM. In particular, _prop1 will be instantiated with the proposition from the *Inform* action and _rule will be constrained to a rule that S1 might think suggests that _prop2 and _prop1 are inconsistent.

However, it is not enough that these instantiations plausibly satisfy the applicability conditions for the *Express-Doubt* action. For example, consider the following dialogue:

---

[5]Some utterances, such as (1) and (6), do not relate to previous dialogue on the discourse level.

(9) S1: *Who is teaching CS360?*
(10) S2: *Dr. Smith.*
(11) S1: *Isn't Dr. Smith the professor who won the teaching award last year?*

While it is at least plausible that, in the mind of the speaker, there is a rule which makes winning a teaching award inconsistent with teaching CS360, interpreting (11) as expressing doubt at (10) seems incorrect. Thus, to prevent such erroneous interpretations, we contend that evidence is needed to recognize discourse actions that capture the relationship between a new utterance and the existing dialogue context.

In our recognition algorithm, evidence may take two forms: 1) world knowledge indicating that the applicability conditions for an e-action are satisfied, and 2) linguistic evidence from clue words suggesting a particular discourse action. If the system has evidence that the applicability conditions of an e-action are satisfied, then the system will use the knowledge as evidence that this may be the discourse action that the speaker is pursuing. On the other hand, if there is sufficient linguistic knowledge suggesting a particular discourse action, then these applicability conditions should be attributed to the speaker, as long as they are plausible (i.e., if there is nothing in the system's model of the speaker's beliefs to suggest that the applicability conditions are not satisfied). So, if the clue word *but* is used, then a non-acceptance discourse action such as expressing doubt should be easier to recognize (i.e., should require less evidence that the applicability conditions hold) than if the clue word is not present.

For example, consider the following, in which there are no clue words in (14_a) and (14_b), but (14_c) contains the clue word *but*.

(12)   S1: *Who is teaching CS360?*
(13)   S2: *Dr. Smith.*
(14_a) S1: *Isn't Dr. Smith on sabbatical next year?*
(14_b) S1: *Isn't Dr. Smith the professor who won the teaching award last year?*
(14_c) S1: *But isn't Dr. Smith the professor who won the teaching award last year?*

Chaining from (14_a) could produce an inference path containing an *Express-Doubt* discourse action. The surface form of (14_a) establishes that S1 has an uncertain belief that Dr. Smith is on sabbatical, the first applicability condition for an *Express-Doubt* action (see Figure 2).[6] The effect of the question/answer pair in (12) and (13) is that S1 believes that S2 thinks that Dr. Smith teaches CS360, the second applicability condition in an *Express-Doubt* action. The stereotype module contains the belief that professors on sabbatical do not teach, so the system can ascribe to S1 the following belief: $\forall x, y \ (course(y) \land professor(x) \land on\text{-}sabbatical(x)) \rightarrow \neg teaches(x, y)$. This belief satisfies the third applicability condition (the belief about a rule), and this belief, along with the belief that Dr. Smith is on sabbatical, implies that Dr. Smith is not teaching CS360, the fourth applicability condition. Finally, a check of the DM indicates that the proposition

---

[6] The first applicability condition is actually an uncertain belief. [LC92] describes our formal belief model and how belief strengths are represented in recipes and in our belief model.

that Dr. Smith teaches CS360 is in focus, the last applicability condition. Therefore, since there is evidence for the applicability conditions of the *Express-Doubt* action, and since focusing heuristics suggest that this is a coherent discourse action (although not the most preferred), (14_a) is recognized as expressing doubt at S2's answer, that Dr. Smith is teaching CS360.

If the dialogue included (14_b) or (14_c) instead of (14_a), some of the same evidence would exist. In both (14_b) and (14_c), the system believes 1) that S1 believes that Dr. Smith won the teaching award last year (though S1 is not sure of this), 2) that S1 believes that S2 thinks that Dr. Smith is teaching CS360, and 3) that the proposition that Dr. Smith teaches CS360 is in focus. Thus, some of the applicability conditions for expressing doubt at Dr. Smith teaching CS360 hold. However, the system has no knowledge for the crucial implication that determines how this utterance relates to the preceding dialogue; the system has no evidence that S1 believes that winning a teaching award implies that Dr. Smith is not teaching CS360. Therefore, (14_b) is not interpreted as an expression of doubt at the response in (13), and other discourse acts are considered. The presence of the clue word in (14_c), however, strongly suggests an *Express-Doubt* discourse act and thus less evidence is needed to recognize it; that is, the system does not need explicit evidence that S1 holds the requisite beliefs but only needs to be able to plausibly ascribe them to S1. Therefore, since our model can plausibly ascribe to S1 belief in the implication that Dr. Smith winning a teaching award implies that Dr. Smith is not teaching CS360, the system will recognize (14_a) and (14_c) as expressions of doubt, using evidence from linguistic knowledge for (14_c) and from world knowledge for (14_a).

Thus, we want our system to use linguistic knowledge when present, world knowledge when present, and both when possible. Our algorithm for processing is the following: from the semantic representation of the utterance, infer sequences of actions (inference paths) using plan inference rules. If the applicability conditions for any of these actions are implausible, reject the inference. For actions which are not e-actions, tentatively ascribe the beliefs in the applicability conditions. For actions that are e-actions, determine how much evidence is available for the action. If there is more than one e-action for which there is evidence from both linguistic and world knowledge, then choose the inference path closest to the focus of attention for which there is multiple evidence. If linguistic or world evidence is available for more than one e-action, then choose the inference path closest to the focus of attention with this single supporting piece of evidence. If there is no evidence for any e-action, then choose the inference path which contains no e-actions and is closest to the focus of attention.

5 Completing the Example

We return briefly to the dialogue given in Section 3 to illustrate how our process model uses the above algorithm to recognize complex communicative actions such as expressing doubt as well as implicit acceptance of previous utterances.

Utterances (1) and (2) were discussed earlier,

Domain Level
Take-Course

Problem-Solving Level
Build-Plan
Instantiate-Vars   Instantiate-Vars
Instantiate-Single-Var   Instantiate-Single-Var

----▶ = Enable Arc
——▶ = Subaction Arc

Discourse Level
Obtain-Info-Ref   Obtain-Info-Ref
Ask-Ref   Answer-Ref
                 Inform   Ask-Ref
            Address-Believability
            Address-Unacceptance
Request   Express-Doubt   Resolve-Conflict   Request
     Tell   Convey-   Address-
             Uncertain-   Uncertain-Belief
             Belief   Inform-If
                    Tell-If   Address-Believability
Surface-   Explain-Claim   Surface-
WH-Question Surface-Inform Surface-Neg Surface-Inform   Inform   WH-Question
   (1)   (2)   YN-Question   Yes   Tell   (6)
                                (3)   (4)   Surface-Inform
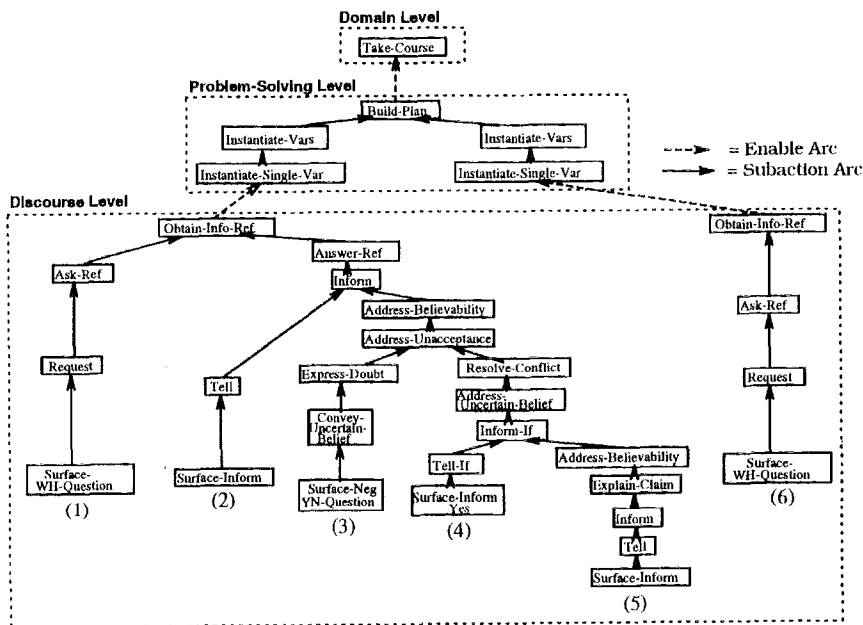   (5)

Figure 3: Dialogue Model for Example

and the DM for these utterances is given in Figure 1. The semantic representation of (3) is: Surface-Neg-YN-Question(S1, S2, undergrad-course(CS360)).

The surface form of the utterance suggests that S1 *thinks* that CS360 is an undergrad course, but is not certain of it. This belief is tentatively entered into the system's model of S1's beliefs. Plan inference rules suggest that the surface request is an action in a *Convey-Uncertain-Belief* action. The linguistic clue *but* suggests that S1 is objecting to some previous utterance of S2's. One of the possible inferences from the *Convey-Uncertain-Belief* action is an *Express-Doubt* action. However, since this is an e-action, evidence is necessary to infer this action. The linguistic clue provides one piece of evidence to support this inference, but the system also looks for evidence from world knowledge. It therefore checks to see if it is plausible that S1's belief about CS360 being an undergrad course might imply that Dr. Smith is not teaching CS360. Although the system has no such belief explicitly represented in its model of S1, there is also no evidence to suggest that S1 does *not* believe that this implication might hold. Since there is no other e-action for which there is evidence and since the applicability conditions for an *Express-Doubt* action are plausible, the system infers that the *Convey-Uncertain-Belief* action may be an action in an *Express-Doubt* action which may be an action in an *Address-Unacceptance* discourse action which may be an action in an *Address-Believability* discourse action which may be an action in an *Inform* action. Focusing heuristics suggest that this Inform is the same *Inform* that (2) is pursuing. Thus, (3) is interpreted as not

accepting (2) by expressing doubt at it.[7]

Inferencing for the remainder of the dialogue is similar to the first three utterances. The inference paths which result from utterances (4) and (5) are shown in Figure 3 above the appropriate numbers.

Finally, although the system initially attempts to closely relate utterance (6) to other utterances at the discourse level, there is no evidence for any e-action that might link this utterance to the existing context on the discourse level. There are no clue words to suggest a relationship, and there are no e-actions for which there is evidence that the applicability conditions hold. Therefore, a completely new discourse action of obtaining information is inferred.[8] This initiation of a new discourse action indicates implicit acceptance of the previous discourse actions since if S1 did not accept S2's answer, S1 would be required to indicate non-acceptance [LC92]. The DM for the entire dialogue is shown in Figure 3 (for space reasons, only the action names are shown). Thus our model recognizes both acceptance and non-acceptance of communicated

---

[7]Although not discussed, there must also be an e-action that relates (2) to the DM. This action is the *Answer-Ref*; evidence for the *Answer-Ref* action is from world knowledge, which indicates that the applicability conditions for this *Answer-Ref* action hold. Inferencing of (2) is then similar to that of (3).

[8]Further discussion may determine that S1 is trying to continue the negotiation dialogue; however, that has not been communicated by this utterance. We are investigating how our system might modify the DM that it has built if it discovers later that the structure built previously is incorrect.

propositions, including acceptance after negotiation of conflicting beliefs.

This example has illustrated: 1) how the structure of a discourse is identified in our three level model, how our model recognizes the relationship of the current utterance to the existing context and to other utterances, and how the tripartite structure produces a richer model of discourse structure than previous models; 2) how beliefs are communicated, recognized, and used in the identification of discourse actions and discourse structure; and 3) how our process model uses linguistic, world, and contextual knowledge together in order to recognize acceptance and non-acceptance of communicated propositions.

6   Conclusions and Future Work

Our plan-based model of dialogue incorporates world, linguistic, and contextual knowledge sources into the recognition of communicative actions. Linguistic knowledge suggests certain discourse acts, a speaker's beliefs, and the strength of those beliefs; contextual knowledge suggests the most coherent continuations of the dialogue; and world knowledge provides evidence that the applicability conditions hold for those discourse acts that identify the relationship of the current utterance to the discourse as a whole. By combining these different knowledge sources, we are able to recognize complex discourse acts such as expressing doubt, to identify the relationship of utterances to one another, and to capture the rich structure of task-oriented dialogue.

Grosz and Sidner [GS86] claim that a robust model of understanding must use constraint satisfaction to interpret utterances; that is, when evidence is available from one source, less evidence is needed from other sources. We have partially included their suggestion by using world and linguistic knowledge when contextual knowledge is not sufficient to infer actions for which there must be some evidence. However, we would like to expand our notion of partial evidence to allow evidence from the three knowledge sources to be represented in terms of degree: thus, when world knowledge is overwhelmingly strong, no other knowledge is needed, but when it is very weak, knowledge from other sources will be needed to support the inferences not allowed by the weak world knowledge alone.

References

[All79]   James F. Allen. *A Plan-Based Approach to Speech Act Recognition*. PhD thesis, University of Toronto, Toronto, Ontario, Canada, 1979.

[Car87]   Sandra Carberry. Pragmatic Modeling: Toward a Robust Natural Language Interface. *Computational Intelligence*, 3:117–136, 1987.

[Car89]   Sandra Carberry. A Pragmatics-Based Approach to Ellipsis Resolution. *Computational Linguistics*, 15(2):75–96, 1989.

[Gri75]   Joseph E. Grimes. *The Thread of Discourse*. Mouton, The Hague, Paris, 1975.

[GS86]   Barbara Grosz and Candace Sidner. Attention, Intention, and the Structure of Discourse. *Computational Linguistics*, 12(3):175–204, 1986.

[Hin89]   Elizabeth Hinkelman. Two Constraints on Speech Act Ambiguity. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pages 212–219, Vancouver, Canada, 1989.

[LA87]   Diane Litman and James Allen. A Plan Recognition Model for Subdialogues in Conversation. *Cognitive Science*, 11:163–200, 1987.

[LC91]   Lynn Lambert and Sandra Carberry. A Tripartite Plan-based Model of Dialogue. In *Proceedings of the 29th Annual Meeting of the ACL*, pages 47–54, Berkeley, CA, June 1991.

[LC92]   Lynn Lambert and Sandra Carberry. Modeling Negotiation Dialogues. In *Proceedings of the 30th Annual Meeting of the ACL*, Newark, DE, June 1992. To appear.

[MT83]   William C. Mann and Sandra A. Thompson. Relational Propositions in Discourse. Technical Report ISI/RR-83-115, ISI/USC, November 1983.

[Pol86]   Livia Polanyi. The Linguistics Discourse Model: Towards a Formal Theory of Discourse Structure. Technical Report 6409, Bolt Beranek and Newman Laboratories Inc., Cambridge, Massachusetts, 1986.

[Pol90]   Martha Pollack. Plans as Complex Mental Attitudes. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*, pages 77–104. MIT Press, 1990.

[Ram91]   Lance A. Ramshaw. A Three-Level Model for Plan Exploration. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 36–46, Berkeley, California, 1991.

[Rei78]   Rachel Reichman. Conversational Coherency. *Cognitive Science*, 2:283–327, 1978.

[Sid81]   Candace L. Sidner. Focusing for Interpretation of Pronouns. *American Journal of Computational Linuistics*, 7(4):217–231, 1981.

[Sid85]   Candace L. Sidner. Plan Parsing for Intended Response Recognition in Discourse. *Computational Intelligence*, 1:1–10, 1985.

[Wil81]   Robert Wilensky. Meta-Planning: Representing and Using Knowledge About Planning in Problem Solving and Natural Language Understanding. *Cognitive Science*, 5:197–233, 1981.