MODELING THE USER'S PLANS AND GOALS

Sandra Carberry

Department of Computer and Information Sciences
University of Delaware
Newark, DE 19716

This work is an ongoing research effort aimed both at developing techniques for inferring and constructing a user model from an information-seeking dialog and at identifying strategies for applying this model to enhance robust communication. One of the most important components of a user model is a representation of the system's beliefs about the underlying task-related plan motivating an information-seeker's queries. These beliefs can be used to interpret subsequent utterances and produce useful responses. This paper describes the IREPS system, emphasizing its dynamic construction of the task-related plan motivating the information-seeker's queries and the application of this component of a user model to handling utterances that violate the pragmatic rules of the system's world model. By reasoning on a model of the user's plans and goals, the system often can deduce the intended meaning of faulty utterances and allow the dialogue to continue without interruption. Some limitations of current plan inference systems are discussed. It is suggested that the problem of detecting and recovering from discrepancies between the system's model of the user's plan and the actual plan under construction by the user requires an enriched model that differentiates among its components on the basis of the support the system accords each component as a correct and intended part of the user's plan.

1 INTRODUCTION

Ideally, a natural language system's responses should contain exactly the information that will be most helpful to the user. But since not all users are alike, achieving such behavior requires that the system have a model of the particular user with whom it is currently interacting. One way of constructing this model is to query the user at the start of the interaction, as was done in the GRUNDY system (Rich 1979). But querying the user may fail to provide an accurate and adequate characterization. Or extensive questioning of the user may be inappropriate if one of the system's goals is that its dialog with the user resemble naturally occurring information-seeking dialogs. In such cases, the system may be able to use the information exchanged during the dialog and its knowledge of the domain to hypothesize a model of the user, and dynamically adjust and expand the model as the dialog progresses.

One of the most important components of a user model is the representation of the system's beliefs about the user's plans and goals. As demonstrated by Cohen, Perrault, and Allen (1981), users of question answering systems "expect more than just answers to isolated questions. They expect to engage in a conversation whose coherence is manifested in the interdependence of their often unstated plans and goals with those of the system."

We are interested in a class of information-seeking dialogs in which the information-seeker is attempting to construct a plan for a task. The task is not being performed during the system's interaction with the user, as is the case in apprentice-expert dialogs, but instead is being constructed for future execution. In some cases, only a partial plan will be constructed, with further details filled in later. For example, a freshman accessing an advisement system may only construct part of his plan for earning a degree, leaving other aspects of the plan to be fleshed out in subsequent years. These dialogs are typical of a large percentage of interactions with database management systems, decision support systems, and expert systems. Typical tasks include expanding a company's product line, purchasing a home, or pursuing a degree at a university. One aspect of our research has been to develop a strategy for dynamically constructing a model of an information-seeker's underlying task-related plan from an ongoing dialog and for tracking his focus of attention in this plan. Since the system's beliefs about the user's plans and

goals provide a context for understanding subsequent dialog, we will refer to this part of a user model as a *context model*.

Our analysis of naturally occurring dialog indicates that humans understand many utterances that would appear imperfect or incomplete to current natural language systems. For example, a speaker may inadvertently use incorrect or ambiguous terminology in constructing the language representation of his intended query, or he may shortcut its complete specification. But if a system's communication is to be regarded as natural, it must be able to handle the full spectrum of utterances that humans appear to understand with relative ease.

Part of this research has been concerned with how a system can reason on its context model to remedy many of a user's faulty utterances and allow the dialog to continue without interruption. We have developed a method, based on Grice's theory of meaning and maxim of relevance, for handling queries that do not conform to the system's model of the world and are therefore regarded as ill-formed. This method uses the task-related plan inferred for an information-seeker to suggest variants of an ill-formed utterance that might represent the information-seeker's intentions or at least satisfy his perceived needs. We have also developed a method for understanding intersentential elliptical utterances that occur during the course of an information-seeking dialogue. Our strategy uses discourse expectations and our model of the speaker's plan to identify the discourse goal that he is pursuing via an elliptical fragment and to interpret his elliptical utterance relative to his task-related plan. Our work on understanding ellipsis is presented in Carberry (1985).

Section 2 of this paper briefly reviews related work in plan recognition, and Section 3 presents our strategy for dynamically inferring a model of the user's plan from an ongoing dialog. Section 4 describes our mechanism for reasoning on this context model to understand a class of utterances that is problematic for current natural language systems. These ideas have been implemented and tested in the IREPS system (Intelligent REsPonse System). This is an ongoing research effort whose objective is a robust natural language interface to information systems. The component that infers the user's underlying task-related plan is called TRACK. To move from one domain to another, only the corpus of domain-dependent plans and goals must be reconstructed; the heuristics and processing strategies are completely transportable.

Section 5 describes our current research on relaxing restrictive assumptions present in previous work on plan recognition and developing a more robust plan inference framework. We propose a four-phase approach to handling the problem of possible disparity between the system's context model and the actual plan under construction by the user, and argue for an enriched context model that differentiates among its components according to the support it accords each component as a correct and intended part of the user's plan.

Throughout this paper, the information-seeker and information-provider will be referred to as *IS* and *IP*, respectively.

2 RELATED WORK ON PLAN RECOGNITION IN DIALOG

Early work in dialog understanding concentrated on apprentice-expert dialogs, during which an expert guided an apprentice in performing a task. Grosz (1977) formulated heuristics for recognizing shifts in focus of attention in the task structure and presented a strategy that used the knowledge currently focused on by the dialog participants to identify the referents of definite noun phrases appearing in an utterance. Robinson (Robinson et al. 1980, Robinson 1981) extended Grosz's work and developed a process model for determining the referents of verb phrases, such as variants of *do* in the utterance

"I've done it."

However, in apprentice-expert dialogs, the overall task that the apprentice is attempting to perform is known at the outset of the dialog, and the ordering of actions and subtasks in the plan being executed strongly influences the dialog between expert and apprentice. This differs from the kind of information-seeking dialogs that we are investigating, in which the information-seeker is attempting to construct a plan for a task that will be executed at a later time. In such dialogs, the information-seeker's utterances are not tightly constrained by the order in which actions in the task will eventually be executed. For example, in an information-seeking dialog between a client and a travel agent, the client may first plan hotel accommodations and theater attendance in New York before inquiring about ways to reach New York, even though *travel to New York* will occur before *attend a New York theater* in a temporal ordering of actions in the resultant plan.

Allen (Allen et al. 1980, Perrault et al. 1980) inferred the goal underlying a speaker's utterance in the context of an information agent in a train setting. This inferred goal was used to account for extra helpful information included in the agent's response, and the inference path connecting the speaker's utterance and the inferred goal was used to interpret indirect speech acts. However, Allen's domain was very restricted; the only domain goals were *meeting a train* and *boarding a train*, each of which could be accomplished by a few primitive steps, and his system was primarily concerned with utterances that might occur at the outset of a dialog. In more complex domains, the information-seeker's complete plan will consist of a hierarchy of subplans and subgoals that accomplish his overall goal. Such a complete plan is not immediately evident from a single utterance, and

Learn-Material(_agent:&PERSON, _sect:&SECTIONS, _syl:&SYLLABI)
   Plan-Body:
         Learn-From-Person(_agent:&PERSON, _sect:&SECTIONS, _fac:&FACULTY)
            where Teaches(_fac:&FACULTY, _sect:&SECTIONS)
         Learn-From-Text(_agent:&PERSON, _txt:&TEXTS)
            where Uses(_sect:&SECTIONS, _txt:&TEXTS)
   Effects:
         Learn-Material(_agent:&PERSON, _sect:&SECTIONS, _syl:&SYLLABI)

Figure 1. Plan For Learning Course Material.

individual utterances must be related to one another to build the user's plan as the dialog progresses.

Sidner (1983, 1985) and Litman (1986) developed enhanced models of plan inference. However, both were concerned with dialogs that were initiated in order to begin or continue execution of an underlying task (display of structures on a graphics terminal and meeting/boarding a train), and the dialogs were therefore constrained by the order in which individual actions in the task had to be executed. In addition, Sidner investigated how discourse markers aid in recognizing the speaker's intent, and Litman studied a meta-plan framework for task dialogs.

Pollack (1986) has recently proposed that plans be viewed as mental phenomena. She contends that, in order to comprehend an utterance and relate it to the user's plan, the system must reason about the configuration of beliefs and intentions that it should ascribe to the speaker. This will be discussed further in Section 5.

3 INFERRING AND MODELING THE TASK-RELATED PLAN

In order to reason about what the user wants to accomplish, the system must have knowledge about the goals that a user might pursue in a domain and plans for accomplishing these goals. We view a plan as the means by which an agent can accomplish a non-primitive task-related goal. Using an extension of the STRIPS formalism (Fikes et al. 1971), we represent a plan as a structure containing applicability conditions, preconditions, a plan body, and effects.

Applicability conditions and preconditions both represent conditions that must exist before a plan can be executed. However, an agent can plan to satisfy preconditions, whereas it is generally anomalous to plan to satisfy applicability conditions; the latter determine whether it is reasonable to even consider a particular plan for achieving a desired goal. For example, suppose an agent wants to vacation on a particular island. If the island has an airport and the agent has money for a ticket, then the agent can plan to fly there. But the requirements that the island have an airport and that the agent have money for a ticket are intuitively different. If the agent does not have enough money for a ticket, he can plan to try and satisfy this requirement; but if the

island does not have an airport, it is unreasonable for the agent to arrange for an airport to be built on the island so that he can fly there for a vacation.

Of course, agents sometimes do unreasonable acts. If the agent in the above case is very wealthy, is adamant about vacationing on this particular island, and abhors boats, he may build an airport on the island and charter a plane to fly him there. Our plans are intended to represent normal plans that an agent might be expected to pursue, and the distinction between preconditions and applicability conditions is useful in preventing consideration of plans that would occur only in exceptional circumstances. How exceptional plans should be incorporated into a plan recognition system is an area for future work.

Wilkins used preconditions similar to our applicability conditions in representing operators in the SIPE system (Wilkins 1984). What we call a precondition, Wilkins incorporated into the set of actions and goals comprising an operator. His reasons for having unplannable preconditions in his representation scheme were both to capture the appropriateness of applying an operator in a given situation and to connect different levels of detail in a hierarchical planner. A proposition, that at one level of abstraction was part of the specification of how an operator was to be performed, might appear at a lower level of abstraction as a precondition of an operator, indicating that further planning for the lower level operator is inappropriate unless the proposition is already satisfied. But mixing standard preconditions (conditions that must exist before an operator can be performed, but which can be planned for) with the set of goals and actions that constitute performing an operator fails to capture the intuitive difference between the two. For this reason, our representation scheme distinguishes among applicability conditions, preconditions that can be planned for, and how one goes about performing an action.

The plan body contains a conjunction of subgoals, and the effects represent the results of successfully executing the plan. Arguments in plans are either constants, represented as uppercase strings, or typed variables, represented as lowercase strings preceded by the character "_" and followed by the characters ":&" and

an uppercase string giving the variable's type. Figure 1 presents a sample plan used by TRACK. Its plan body states that in order to learn the material of a section of a course, an agent must both learn from the person teaching the section and study the text used in the section. TRACK's plans are hierarchical, since many of the subgoals in the bodies of plans and many preconditions are non-primitive and therefore have associated plans which may be substituted for them. Thus a plan can be expanded to any desired degree of detail by replacing non-primitive preconditions and subgoals with their associated plans.

At the outset of an information-seeking dialog, the system has little knowledge about the information-seeker's (IS) purpose in requesting information. In most cases, a complete plan for IS cannot be constructed during the first part of a dialog. Instead, potential goals must be inferred from individual utterances and integrated into the overall plan structure, thereby incrementally expanding and instantiating the system's model of IS's plan as the dialog progresses.

Oftentimes there will be many domain goals that a single utterance might address. For example, if a student asks what time Dr. Smith arrives in the morning, he may want to either visit Dr. Smith or call him in his office. Furthermore, even if we can identify a single domain goal addressed by an utterance, there may be many ways in which that goal could be incorporated into an overall plan. For example, if a student asks whether Political Science 210 is offered in the spring, we might infer that the student wants to take Political Science 210. But how should this goal be built into the student's overall plan? Perhaps he is considering taking Political Science 210 in order to satisfy a breadth requirement or perhaps he wants to major or minor in Political Science. So the issue that must be addressed in dynamically inferring IS's underlying task-related plan from an ongoing dialog is the following: *how can we identify which of many candidate goals is the actual goal which IS is addressing with a particular utterance, and how can we determine where this particular goal fits into IS's overall plan?*

Two factors appear to provide a basis for a solution: 1) the organized nature of naturally occurring information-seeking dialogues, as exhibited in dialog transcripts, and 2) the assumption that IS and IP are working cooperatively to help IS achieve his plan construction goals. These two factors produce a structure in information-seeking dialogs. As a result, we can formulate focusing heuristics that specify how individual utterances should be related to the existing dialog context, as represented by the plan inferred for IS and his current focus of attention in that plan.

Thus our approach is the following:
1. hypothesize from an individual utterance a set of domain-dependent candidate focused goals that may represent the information-seeker's focus of attention in the task; and

2. use focusing heuristics to select the candidate focused goal most apropos to the existing dialogue context and incorporate it into the model of the information-seeker's plan.

In some cases, several candidate focused goals may be equally likely, and alternative versions of the context model may need to be built. In a cooperative, coherent dialog in which the information-seeker successfully communicates how his questions relate to what he wants to accomplish, subsequent utterances should enable the system to identify the particular context model that represents the user's plan. However, if the user asks a sequence of questions that have no definite relationship to one another, then we may have a computationally explosive situation. But this behavior violates our assumption of an overall cooperative, coherent dialog.

We have given some preliminary consideration to a more robust process model containing a stack of disjoint contexts with potential relationships to one another. Further utterances could clarify these relationships and permit merger of the disjoint contexts into a single overall context model. Such a strategy would have the advantage of handling disconnected portions of dialogs without the computational explosion that can result from modeling all possible expanded contexts.

3.1 HYPOTHESIZING CANDIDATE FOCUSED GOALS AND PLANS

The first stage of processing analyzes an utterance without considering the preceding dialog. Plan-identification heuristics are used to hypothesize a set of domain-dependent goals and associated plans that might represent that aspect of the task on which IS's attention is currently focused. These heuristics are extensions of inference rules proposed by Allen (Allen et al. 1980). For example, if IS wants to know the values of an argument that cause a proposition to be true, then that proposition or its negation may be relevant to the plan that IS is considering. Therefore any goals whose plans might prompt such a request become *candidate focused goals* and their associated plans become *candidate focused plans*. Thus if IS asks,

> "Who is teaching section 10 of French 112 in the spring of 1988?"

then IS wants to know the values of the argument _fac: &FACULTY such that the proposition

$$\text{Teaches(\_fac:\&FACULTY, FRENCH112-10-SPRING88)}$$

is true. The plan for learning the material of a section of a course (Figure 1) contains the proposition

$$\text{Teaches(\_fac:\&FACULTY, \_sect:\&SECTIONS)}$$

as a constraint on a subgoal in its body. Substituting FRENCH112-10-SPRING88 for _sect:&SECTIONS produces the proposition
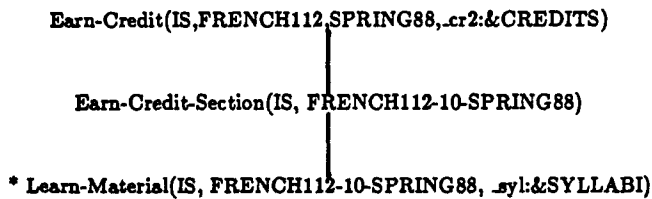
Earn-Credit(IS,FRENCH112,SPRING88,_cr2:&CREDITS)

Earn-Credit-Section(IS, FRENCH112-10-SPRING88)

* Learn-Material(IS, FRENCH112-10-SPRING88, _syl:&SYLLABI)

Figure 2. An Example of a Context Model.

Teaches(_fac:&FACULTY,
FRENCH112-10-SPRING88)

addressed by IS's utterance. Making this substitution throughout the plan in Figure 1 produces a plan for learning the material of section 10 of French 112 in the spring of 1988. Therefore the goal

Learn-Material(IS, FRENCH112-10-SPRING88,
_syl:&SYLLABI)

becomes a candidate focused goal, and the plan produced by substituting FRENCH112-10-SPRING88 for _sect:&SECTIONS in Figure 1 becomes a candidate focused plan. The goal

Learn-From-Person(IS, FRENCH112-10-SPRING88,
_fac:&FACULTY)

where

Teaches(_fac:&FACULTY,
FRENCH112-10-SPRING88)

is the most recently considered subgoal in this candidate focused plan; thus it provides the greatest expectations for future utterances.

3.2 CONTEXT PROCESSING

The second stage relates an utterance to the context established by the preceding dialog. We use a tree structure called a *context model* to represent the task-related plan inferred for IS from the preceding dialog. Each node in this tree represents a goal that the system believes IS is considering achieving and, except for the root, is a descendant of a higher-level goal whose associated plan contains the subgoal represented by the child node. In Figure 2, for example, learning the material for section 10 of French 112 appears as a descendant of the higher-level goal of earning credit in section 10 of French 112, representing the belief that IS is considering how he would go about learning the material of the section as part of a plan for earning credit in the section.

One node in the tree is marked as the current focus of attention and indicates that aspect of the task on which IS's attention is currently centered. The path from the root of the context model to the current focus of attention is called the *active path* and represents the global context, or sequence of progressively lower-level goals that led to the subgoal currently under consider-

ation by IS. In Figure 2, the active path contains the goals

Earn-Credit(IS,FRENCH112,SPRING88,
_cr2:&CREDITS)
Earn-Credit-Section(IS,FRENCH112-10-SPRING88)
Learn-Material(IS, FRENCH112-10-SPRING88,
_syl:&SYLLABI)

with the current focus of IS's attention believed to be

Learn-Material(IS, FRENCH112-10-SPRING88,
_syl:&SYLLABI)

Initially, there is no existing context; each candidate focused goal and its plan become the root of a context model and are marked as current focused goals and plans. If there is only one context model and its root goal appears as part of only one domain-dependent plan, then we have further knowledge about what IS wants to do and can add this higher-level plan as the new root of the context model, with the old root as its child. We continue expanding the tree upward until more than one higher-level plan is possible. For example, if IS's first utterance was

"Who is teaching section 10 of French 112 in the
spring of 1988?"

then, as described in the previous section,

Learn-Material(IS, FRENCH112-10-SPRING88,
_syl:&SYLLABI)

would become a candidate focused goal. This goal and its associated plan would be entered as the root of a context model and be marked as the current focus of attention. In addition, since only the plan for earning credit in section 10 of French 112 contains this goal, and since only the plan for earning credit in French 112 contains the goal of earning credit in a section of French 112, the context model is expanded upward to include these higher-level goals, producing the context model in Figure 2. Thus, only that part of the user's task-related plan that the system believes IS intended it to recognize is built into the context model. Section 5 discusses more robust user modeling, in which default inference rules might be used to expand the system's model of the user's plan, and addresses the problem of detecting and recovering from errors that might be introduced into the model.

As each new utterance occurs, it must be related to the established context. A set of focusing heuristics are used to determine the most likely relationship between one of the hypothesized candidate focused plans and the context model, and to expand the context model to include it. Grosz (1977) introduced the concept of focusing in her work on identifying the referents of definite noun phrases in apprentice-expert dialogs. She noted that the focus of the discourse followed the plan for performing the apprentice's task. Our information-seeking dialogs differ from apprentice-expert dialogs in

that our dialogs are not constrained by the order of execution of the actions in the overall plans. However, we do find structure in the dialogs we are studying, and it is the basis for our focusing heuristics. This structure appears to be caused by two factors.

The first is the organized nature of naturally occurring information-seeking dialogues. Dialogue transcripts indicate that humans generally ask all their questions that are relevant to a plan for one subgoal before moving on to ask questions about a plan for another subgoal of the overall task. One possible explanation for this behavior is that it may require less mental effort than switching back and forth among partially constructed plans for different subgoals.

The second factor producing structure in our dialogs is their cooperative nature. Since the dialogs are cooperative and miscommunication can occur if both dialog participants are not focused on the same subset of knowledge (Grosz 1981), we expect IS to shift topic slowly between consecutive utterances and to adhere to the focusing constraints espoused by McKeown (1985). McKeown expanded on focus rules proposed by Sidner (1981) to explain how speakers should organize their utterances when faced with a choice of topic. In particular, McKeown claims that a speaker should move to a recently introduced topic if he has something further to say about it; otherwise he will have to reintroduce the topic at a later time. Similarly, the speaker should choose to finish discussion of the current topic before switching back to a previous one.

Our focusing heuristics rely on these expectations about possible shifts in focus of attention in IS's underlying task-related plan to identify which candidate focused plan is most apropos to the established dialog context and to determine how it fits into the context model. The following ordered list gives the focusing heuristics' preferences on the relationship between a candidate focused plan and the context model. Each relationship is illustrated under the assumption that the tree shown in Figure 3a is the context model immediately preceding the utterance, with node C (marked by an asterisk) representing the current focused goal/plan and node G representing the most recently considered subgoal in the current focused plan.

1. The candidate focused plan is part of the expansion of a plan for the most recently considered subgoal in the current focused plan; for example, if Figure 3b is an expansion of the context model shown in Figure 3a, then node C1 might represent such a candidate focused plan.

2. The candidate focused plan is part of an expansion of the current focused plan. For example, if Figure 3b is an expansion of the context model shown in Figure 3a, then node C2 might represent such a candidate focused plan (where C2 is part of a plan for the goal at node F, which is in turn part of a plan for the goal at node C).

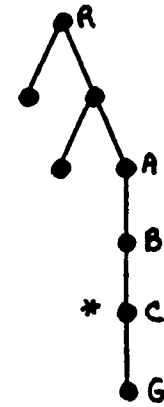3. The candidate focused plan is part of the expan-
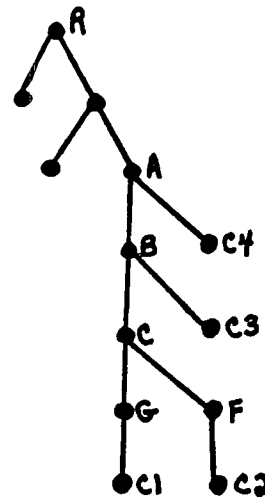

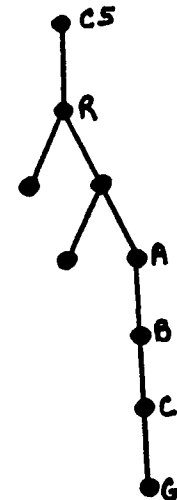Figure 3a. Initial Context Model.

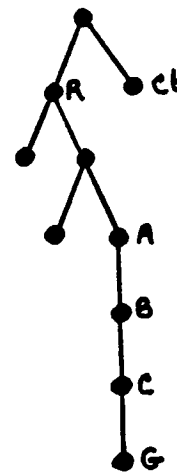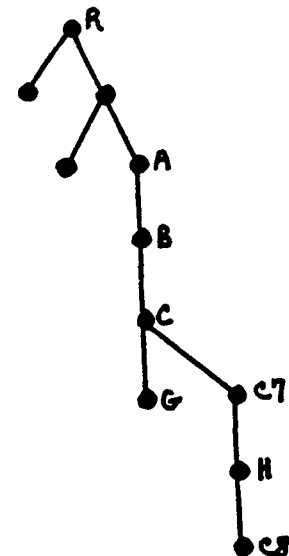Figure 3b.

Figure 3c.

Figure 3d.

Figure 3e.

Figure 3. Examples of Relationships Between Candidate Plans and Context Model.

* Satisfy-Major(IS, BA, CS)   * Satisfy-Major(IS, BS, CS)

Figure 4. Two Context Models.

sion of a plan for a goal along the active path, with preference given to goals that are closest to the current focused goal on the active path. For example, if Figure 3b is an expansion of the context model shown in Figure 3a, then nodes C3 and C4 would both be part of the expansion of a plan for a goal along the active path; but if nodes C3 and C4 both represent candidate focused plans, then node C3 would be preferred, since it appears in an expansion of the plan associated with node B, which is closer to the current focused goal (represented by node C) than is node A.

4. The candidate focused plan is a plan whose expansion contains the goal associated with the root of the context model; Figure 3c illustrates such a relationship, where node C5 represents a candidate focused plan.

5. The candidate focused plan is part of the expansion of a higher-level plan, and this expansion also contains the goal associated with the root of the context model; Figure 3d illustrates such a relationship, where node C6 represents a candidate focused plan.

In applying each rule, we use a breadth-first expansion of plans, so that the resulting shift in focus of attention will be as small as possible. For example, if nodes C7 and C8 in Figure 3e both represented candidate focused goals/plans, the second rule in the above list would prefer C7 to C8, since C7 is closer to the existing focus of attention in the dialog.

3.3 AN EXAMPLE

To illustrate this plan inference process, let us consider a dialog segment containing four utterances by IS. Suppose IS begins with the statement

"I want to major in computer science."

Since IS states that he wants to achieve a goal, majoring in CS, TRACK's plan identification heuristics hypothesize

Satisfy-Major(IS, BA, CS)
and
Satisfy-Major(IS, BS, CS)

as candidate focused goals and their associated plans as candidate focused plans. Since there is no way of choosing between these, two context models would be built, each with one of the candidate focused goal/plan pairs as its root (Figure 4). The resulting current focused plan in each context model is preceded by an asterisk.

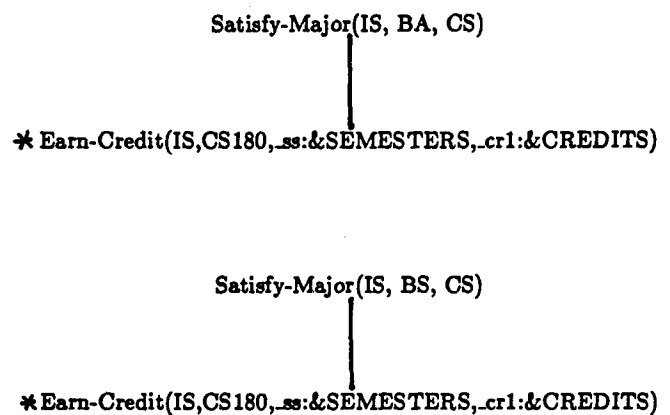Suppose that IS's next utterance is the query

Satisfy-Major(IS, BA, CS)

✹ Earn-Credit(IS,CS180,_ss:&SEMESTERS,_cr1:&CREDITS)

Satisfy-Major(IS, BS, CS)

✹ Earn-Credit(IS,CS180,_ss:&SEMESTERS,_cr1:&CREDITS)

Figure 5. Context Models After Two Utterances.

"What are the prerequisites for taking CS180?"

Since IS wants to know the preconditions for the plan associated with the goal of taking CS180 (the introductory course for majors and minors in computer science), TRACK hypothesizes

Earn-Credit(IS, CS180, _ss:&SEMESTERS, _cr1:&CREDITS)

and its associated plan as the candidate focused goal/plan pair. The focusing heuristics must determine how this relates to the preceding dialog, as represented by the context model. The strongest expectation is that IS is continuing with some aspect of the current focused plan; since the preceding utterance did not address any particular goal in this plan, there is no most recently considered subgoal. Since taking CS180 appears in an expansion of the plans for majoring in computer science, TRACK expands the context models as shown in Figure 5 and marks the plan for earning credit in CS180 as the new current focus of attention.

Suppose that IS's next query is,

"What courses must I take in order to satisfy the foreign language requirement?"

Since IS is asking about the argument (courses) of a subgoal (taking courses) that is part of a plan for achieving a second goal, TRACK hypothesizes the second goal and its associated plan

Satisfy-Language-Req(IS)

as the candidate focused goal/plan pair. The focusing heuristics must now determine how this relates to the preceding dialog, as represented in the context model. The strongest expectation is that IS will continue with some aspect of the current focused plan. However, the candidate focused plan does not appear in an expansion of the current focused plan, indicating that IS has shifted focus to another aspect of the overall task. In fact, none of the first four focusing heuristics find a relationship between the candidate focused plan and the
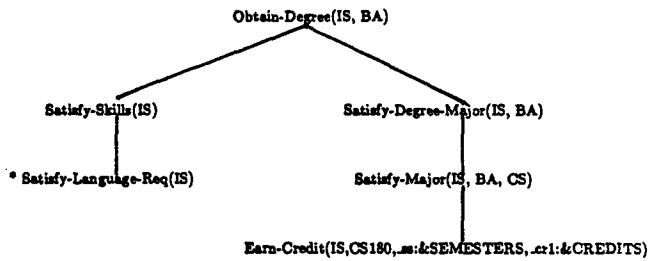
Obtain-Degree(IS, BA)

Satisfy-Skills(IS)   Satisfy-Degree-Major(IS, BA)

* Satisfy-Language-Req(IS)   Satisfy-Major(IS, BA, CS)

Earn-Credit(IS,CS180,_m:&SEMESTERS,_cr1:&CREDITS)

Figure 6. Context Model After Three Utterances.

context model. However, the last focusing heuristic finds that there is a goal,

Obtain-Degree(IS, BA)

whose associated plan can be expanded to include both the candidate focused plan and the context model whose root is Satisfy-Major(IS, BA, CS), indicating that IS has shifted his attention to another subtask (satisfying the foreign language requirement) of a higher-level plan (obtaining a bachelor of arts degree), of which the old current focused plan (obtaining a computer science major) is also a part. Therefore this goal becomes the root of a new context model, as shown in Figure 6; Satisfy-Language-Req(IS) is marked as the new current focus of attention, as indicated by the asterisk preceding it. The other previous context model, whose root was Satisfy-Major(IS, BS, CS), is discarded, indicating that IS's third utterance has led us to deduce that he wants to pursue a bachelor of arts degree. Note that our plan inference process makes what Pollack (1987) terms the *appropriate query assumption*—namely, that IS does not ask queries that are inappropriate to his intended goal. This aspect of our plan inference process will be discussed further in Section 5, where we discuss a more robust plan recognition paradigm.

Suppose that IS's next query is,

"Who is teaching section 10 of French 112 in the spring of 1988?"

As described earlier, since IS is asking about the teacher of a particular section of a course, he may be considering the subgoal of learning from that teacher; this subgoal appears in a plan for learning the material of a course, and therefore TRACK hypothesizes the plan associated with the goal

Learn-Material(IS, FRENCH112-10-SPRING88,
_syl:&SYLLABI)

as one of the candidate focused plans. The focusing heuristics find that this candidate focused plan appears in an expansion of the most recently considered subgoal (taking courses) in the current focused plan, and there-

fore it is selected as the new focus of attention and the context model is expanded to include it (Figure 7).

In this manner, our plan inference process dynamically infers from an ongoing dialog the underlying task-related plan motivating an information-seeker's queries and tracks his focus of attention in this plan structure.

4 APPLICATION OF CONTEXT MODELS

The context model is one component of a comprehensive user model, representing the system's acquired beliefs about the plan an information-seeker is trying to construct. The possible expansions of this plan provide expectations about information that IS might want, and these expectations can often be used to repair and disambiguate IS's subsequent utterances. We have developed strategies that use our context model to handle two forms of problematic input: pragmatically ill-formed utterances and intersentential ellipsis. This section describes our approach to the first of these; our framework for handling ellipsis is described in Carberry (1985).

4.1 PRAGMATIC ILL-FORMEDNESS

An utterance can be syntactically and semantically well formed, yet violate the structural properties of the listener's world model. This is not to say that the speaker necessarily holds an incorrect view of the world, or even one that differs from the listener's view, but only that the semantic representation of the speaker's utterance does not conform to the listener's world model. We shall say that such an utterance is *pragmatically ill-formed*.

Consider, for example, the query

IS: "What is the area of the special weapons magazine of the Alamo?"

that appears in a dialog transcript of an information-seeker attempting to load cargo onto ships using the REL natural language interface (Thompson 1980). A semantic representation of this query will contain the proposition

Area(SPECIAL-WEAPONS-MAG, _areaval:&SQ-FT)

The system was unable to understand this query, since its semantic representation erroneously presumed that storage locations had an area attribute in the associated data base. If a human information-provider had a similar problem in understanding the utterance, or considered the meaning of "area" ambiguous, he might be able to use the context established by the preceding dialog to identify what the information-seeker really wanted to know. For example, if IS's goal was to load cargo of the appropriate type into the various cargo holds, then he probably wanted to know the *remaining* capacity of the Special Weapons Magazine. On the other hand, if his goal was to assign ships to routes in order to handle the
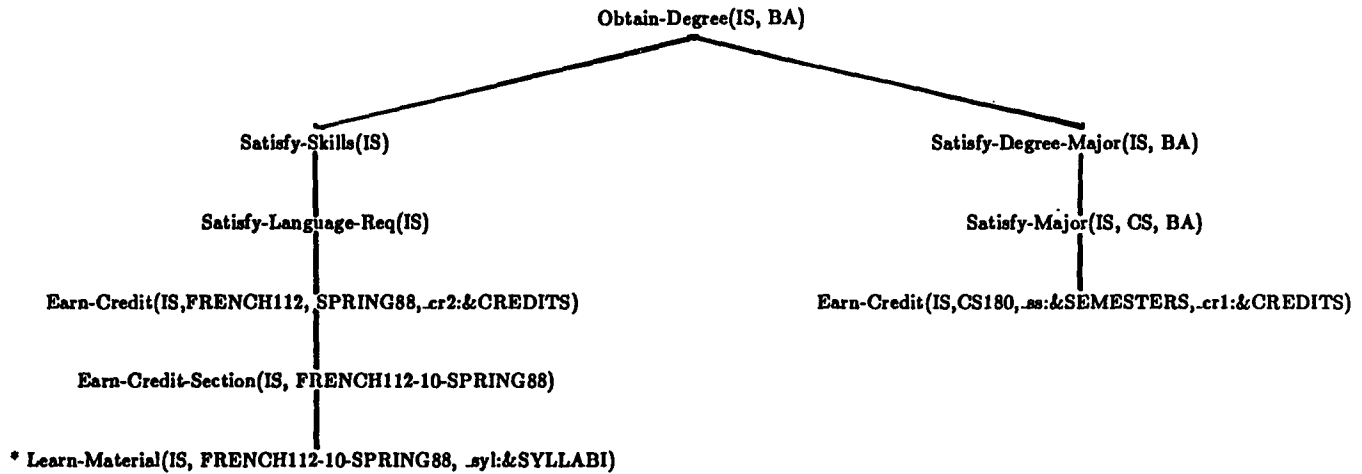
Obtain-Degree(IS, BA)

Satisfy-Skills(IS)   Satisfy-Degree-Major(IS, BA)

Satisfy-Language-Req(IS)   Satisfy-Major(IS, CS, BA)

Earn-Credit(IS,FRENCH112, SPRING88,_cr2:&CREDITS)   Earn-Credit(IS,CS180,_ss:&SEMESTERS,_cr1:&CREDITS)

Earn-Credit-Section(IS, FRENCH112-10-SPRING88)

* Learn-Material(IS, FRENCH112-10-SPRING88, _syl:&SYLLABI)

Figure 7. Context Model After Four Utterances.

expected cargo shipping requirements, then IS probably wanted to know the *total* capacity of the Special Weapons Magazine. Similarly, if his goal was to assign workers to fill the storage holds, with one worker assigned to handle all cargo holds located in the same section of the ship, then IS probably wanted to know the location of the Special Weapons Magazine.

Another example of a pragmatically ill-formed query illustrates the *missing joins* problem.

IS: "Who is teaching section 10 of French 112 in the spring of 1988?"
IP: "Dr. Walker."
IS: "When's Mitchel meet?"

A semantic representation of the last query contains the proposition

$$\text{Meeting-Time(MITCHEL,}$$
$$\_\text{tme:\&MEETING-TIMES)}$$

Suppose that in the system's world model, faculty teach sections of courses, chair committees, and present colloquia, and each of these has a scheduled meeting time, but there is no direct relationship between faculty and times. Then the above query will appear pragmatically ill-formed. Although this utterance might be an abbreviated version of any of the queries

"When does the section of French 112 taught by Dr. Mitchel meet?"
"When does the committee chaired by Dr. Mitchel meet?"
"When does the colloquium given by Dr. Mitchel meet?"

a human information-provider would be likely to recognize from the above dialog that IS wants the meeting time of the section of French 112 taught by Dr. Mitchel, and respond accordingly.

4.2 UNDERSTANDING PRAGMATICALLY ILL-FORMED QUERIES

If a natural language system's communication is to be regarded as natural, the system must be able to handle the full spectrum of utterances that humans understand with relative ease. But our analysis of naturally occurring dialog indicates that human listeners understand many utterances that would appear pragmatically ill-formed to current natural language systems. A number of researchers have investigated the problem of handling pragmatically ill-formed queries (Sowa 1976, Chang 1978, Mays 1980, Kaplan 1982), but their strategies were deficient in that they considered the queries in isolation, without using a model of the preceding dialog to address the speaker's intentions.

Grice's theory of meaning (Grice 1969, Grice 1957) and maxim of relation (Grice 1975) suggest that the listener's beliefs about what the speaker is trying to do should be used to recognize the intent behind an ill-formed query. According to Grice's theory, a listener should believe that the speaker believes the listener can infer the intended meaning of an utterance—otherwise the speaker would not have made it. So given a pragmatically ill-formed query, a cooperative listener should attempt to deduce these intentions. Grice's maxim of relation suggests that the speaker's utterance is relevant to the existing dialog context, so the listener should use this context and the focus of attention immediately prior to the problematic utterance to attempt to deduce the speaker's intended meaning and enable the dialog to continue without interruption.

Our strategy is based on this theory of meaning and intention. It uses the context model to suggest substitutions for the erroneous proposition appearing in the semantic representation of IS's pragmatically ill-formed query, thereby producing semantic representations for one or more revised queries, all of which are apropos to what IS is trying to accomplish. If more than one

revised query is proposed, then it must be determined whether any of these is significantly more likely than the others to represent the speaker's intentions or satisfy his perceived needs. Two criteria appear appropriate for comparing suggested revised queries. The first is the relevance of the revised query to the current focus of attention in the dialog. Since we have contended that some shifts in focus of attention in the plan structure are more likely than others, it is reasonable to hypothesize that the more expected the shift in focus of attention that would result from a revised query, the more likely is that query to represent the speaker's intentions. The second criteria for comparing suggested revised queries is the similarity of a revised query to the speaker's actual utterance. For example, color has less semantic similarity to area than does remaining capacity. Therefore substituting "color" for "area" in the example query

"What is the area of the special weapons magazine of the Alamo?"

is a more significant alteration of the query than is substituting "remaining capacity" for "area." As a result, the revised query

"What is the color of the special weapons magazine of the Alamo?"

is less similar to the speaker's actual query than is the revised query

"What is the remaining capacity of the special weapons magazine of the Alamo?"

Therefore our pragmatic ill-formedness processor contains a suggestion mechanism and a selection mechanism. The suggestion mechanism proposes revised queries, all of which are relevant to IS's underlying task-related plan, and the selection mechanism uses the criteria of relevance and semantic similarity to select, from among multiple suggestions, the revised query deemed most likely to represent the speaker's intentions or satisfy his perceived needs.

4.3 REPAIR STRATEGY

4.3.1 PROPOSING REVISIONS

The suggestion mechanism uses two sets of substitution heuristics, one for making simple substitutions of a property, relation, function, or object class for that used by the speaker, and a second set for expanding relational paths to handle the missing joins problem.

As an example of a simple substitution, suppose the dialogue preceding the query

"What is the area of the special weapons magazine of the Alamo?"

indicates that IS's current focused goal within his overall plan is to load cargo of the appropriate type into

the various cargo holds. A subgoal in the plan associated with this goal would be

$$Load\text{-}Type\text{-}Cargo(IS,\_item:\&CARGO),$$
$$\_storearea:\&STORAGE\text{-}AREA$$
where
$$Is\text{-}Type(\_item:\&CARGO, \_cartype:\&CARGO\text{-}TYPE)$$
$$Cargo\text{-}Type(\_storearea:\&STORAGE\text{-}AREA,$$
$$\_cartype:\&CARGO\text{-}TYPE)$$

and a plan for this subgoal would contain the precondition

$$Greater(\_remcap:\&CUBIC\text{-}FT,$$
$$\_itemsize:\&CUBIC\text{-}FT)$$
where
$$Volume(\_item:\&CARGO, \_itemsize:\&CUBIC\text{-}FT)$$
$$Remaining\text{-}Capacity(\_storearea:\&STORAGE\text{-}AREA,$$
$$\_remcap:\&CUBIC\text{-}FT)$$

specifying that the storage area must have room for the cargo item. The property substitution heuristic would examine this plan and suggest substituting either of the propositions

$$Cargo\text{-}Type(SPECIAL\text{-}WEAPONS\text{-}MAG,$$
$$\_cartype:\&CARGO\text{-}TYPE)$$

and

$$Remaining\text{-}Capacity(SPECIAL\text{-}WEAPONS\text{-}MAG,$$
$$\_remcap:\&CUBIC\text{-}FT)$$

for the erroneous proposition

$$Area(SPECIAL\text{-}WEAPONS\text{-}MAG, \_areaval:\&SQ\text{-}FT)$$

appearing in the semantic representation of IS's query, producing suggested semantic representations equivalent to the two revised queries

IS: "What is the cargo type of the Special Weapons Magazine of the Alamo?"

IS: "What is the remaining capacity of the Special Weapons Magazine of the Alamo?"

More formally, this heuristic is represented by the following rule:

If IS's proposition erroneously presumes that a member Obj1 of CLASS1 has a property Att1, then replace property Att1 with property Att2 if the following conditions hold:
1. a proposition specifying property Att2 on a member Obj2 of CLASS1 appears in an expansion of IS's context model.
2. Obj1 and Obj2 unify (Either Obj1 in IS's utterance or Obj2 in the plan proposition refers to a general member of CLASS1, or both refer to the same specific member of CLASS1).

In the context of our student advisement dialogs, suppose a student wants to pursue an independent study project; such projects can be directed by full-time faculty but not by faculty who are extension or on

sabbatical. The student might erroneously follow the utterance

"I want to take an independent study project."
with the pragmatically ill-formed query

"What is the classification of Dr. Smith?"

In a university world model, only students have a classification attribute; this attribute can have values such as Arts&Science-1988, Engineering-1989, and Business-1990. Faculty have attributes such as rank, status, age, and salary. Pursuing an independent study project under the direction of Dr. Smith has the precondition that Dr. Smith's status be full-time or part-time. Our substitution mechanism would analyze the plan for taking an independent study course, and the property substitution rule would suggest substituting the proposition

Status(DR.SMITH, _statusval:&STATUSVALUES)

for the erroneous proposition

Classification(DR.SMITH,
_classval:&CLASSVALUES)

appearing in the semantic representation of the student's query, resulting in a suggested revised semantic representation equivalent to the query

"What is the status of Dr. Smith?"

As an example of the second set of heuristics, the path expansion heuristics, consider again the query

"When's Mitchel meet?"

following the dialog that produced the context model shown in Figure 7. As mentioned earlier, the semantic representation of this query contains the erroneous proposition

Meeting-Time(MITCHEL,
_tme:&MEETING-TIMES)

indicating a direct relationship between faculty and times. Our path expansion heuristics will analyze and expand the context model shown in Figure 7 and note that a plan for the goal

Earn-Credit(IS, FRENCH112, SPRING88,
_cr2:&CREDITS)

can include a path containing the sequence of goals shown in Figure 8. One path expansion heuristic notes that the propositions

Teaches(_fac:&FACULTY, _sec1:&SECTIONS)
Is-Meeting-Time(_sec1:&SECTIONS,
_tme:&MEETINGTIMES)

both appear on this path in the expanded plan, and suggests substituting the conjunction of the propositions
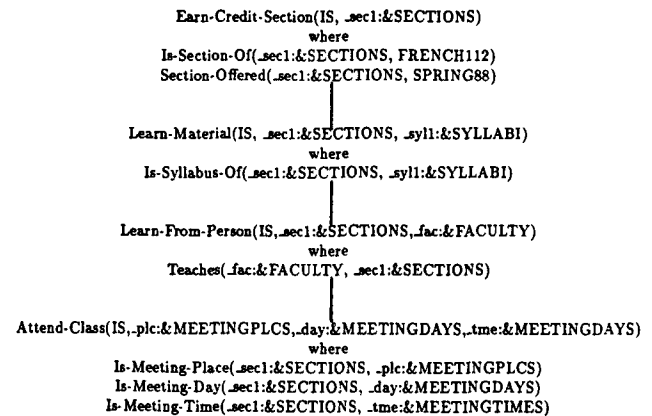
Earn-Credit-Section(IS, _sec1:&SECTIONS)
where
Is-Section-Of(_sec1:&SECTIONS, FRENCH112)
Section-Offered(_sec1:&SECTIONS, SPRING88)

Learn-Material(IS, _sec1:&SECTIONS, _syll:&SYLLABI)
where
Is-Syllabus-Of(_sec1:&SECTIONS, _syll:&SYLLABI)

Learn-From-Person(IS,_sec1:&SECTIONS,_fac:&FACULTY)
where
Teaches(_fac:&FACULTY, _sec1:&SECTIONS)

Attend-Class(IS,_plc:&MEETINGPLCS,_day:&MEETINGDAYS,_tme:&MEETINGDAYS)
where
Is-Meeting-Place(_sec1:&SECTIONS, _plc:&MEETINGPLCS)
Is-Meeting-Day(_sec1:&SECTIONS, _day:&MEETINGDAYS)
Is-Meeting-Time(_sec1:&SECTIONS, _tme:&MEETINGTIMES)

Figure 8.

Teaches(MITCHEL, _sec1:&SECTIONS)
Is-Meeting-Time(_sec1:&SECTIONS,
_tme:&MEETINGTIMES)

for the erroneous proposition appearing in the semantic representation of IS's query, resulting in a revised semantic representation equivalent to the English query

"When do sections taught by Mitchel meet?"

The revised semantic representation no longer violates the system's world model. But it represents an incomplete query, in that it contains an ellipsis. Presumably the speaker wants to know only the sections of French 112 taught by Dr. Mitchel in the spring of 1988, not sections *of any* course taught by Dr. Mitchel during *any* semester. How the context model can be used to interpret elliptical utterances is discussed in Carberry (1985). Although we have only illustrated substituting a conjunction of two propositions for the erroneous proposition in the user's query, the path expansion heuristics can propose expansions of any length.

Five other heuristics and other parts of the user's plan can suggest substitutions in addition to the ones shown in our examples. The important point is that all of the revised semantic representations resulting from these suggestions represent queries that are apropos to the plan that IS is constructing.

4.3.2 SELECTING THE APPROPRIATE REVISION

As mentioned earlier, relevance to the current focus of attention and similarity to the speaker's actual utterance are used to select from among multiple suggestions. We use focusing heuristics, similar to those used for constructing the context model, to measure relevance of a revised query to the current focus of attention in the dialog, and generalization hierarchies for properties, relations, functions, and object classes to measure the semantic similarity of a substituted term and the term that it replaces. In the example

"What is the area of the Special Weapons Magazine of the Alamo?"

both suggested revised queries have approximately the same relevance to the current dialog but, of the two properties cargo type and remaining capacity, remaining capacity is much closer semantically to the property area used by the speaker. Therefore our selection mechanism chooses the semantic representation equivalent to the query

"What is the remaining capacity of the Special Weapons Magazine of the Alamo?"

as the most appropriate interpretation representing IS's needs.

Instead of computing semantic representations for all suggested revised queries and then selecting the best revision, we analyze nodes of the context model in order of decreasing relevance to the existing focus of attention, until a revision meeting an arbitrary level of acceptability is found. This acceptability level initially is set so that only revisions with extremely good evaluations will meet it, and it is steadily relaxed as larger parts of the context model are analyzed. Since one factor used by the evaluation metric is relevance to the existing focus of attention in the dialog, scores for newly suggested revisions will, in most cases, be worse than the scores for revisions suggested much earlier. Thus as more of the context model is analyzed, a revision that previously did not receive a good enough evaluation to terminate processing may now appear more likely to represent the user's intentions. The relaxed acceptability level allows such a revision to be selected as the appropriate interpretation.

This processing mechanism is efficient, since only a small part of the user's expanded plan will usually be analyzed. It also avoids the problem of computational explosion. If processing time exceeds a preset maximum or the acceptability level is relaxed to some preset minimum level of goodness, then the system can terminate its search for an interpretation and is justified in believing that its failure to understand the user's utterance is not unnatural behavior.

4.3.3 COMPARISON TO OTHER STRATEGIES

This approach is superior to previous strategies because it uses a model of the speaker to identify and address his perceived intentions and needs in making an utterance. As such, it not only reasons on the context model to suggest possible interpretations relevant to the user's goals and plans, but it also limits consideration to those interpretations that are reasonable given the established dialog context.

5 IMPROVING PLAN RECOGNITION

Our research has shown how an information-seeker's underlying task-related plan can be dynamically inferred from an ongoing dialog, and how the resulting

context model can be used to achieve better communication. However, the kinds of cooperative information-seeking dialogs handled by current models of plan recognition indicate that four critical assumptions have been made:

1. IS has no misconceptions about the task domain.[1]
2. IS's queries always address aspects of the task within the system's limited knowledge. These systems maintain the *closed world assumption* (Reiter 1978).
3. IS's statements and queries are correct and not unintentionally misleading.
4. The system's inference mechanisms do not introduce errors into the context model.

These assumptions represent unrealistic constraints on real-world dialogs and must be removed. The first assumption, called the *valid plan assumption* by Pollack (1987), limits the kinds of beliefs IS can already have about the domain—namely, it says that IS's knowledge may be incomplete but not erroneous. But IS is interacting with the system because IS does not know enough about the domain to construct his task-related plan by himself. Therefore, since IS is not an expert in the area, it is to be expected that some of his beliefs about the domain may be false, contradicting the first assumption. An implication of the valid plan assumption is what Pollack terms the *appropriate query assumption*—namely, that IS knows enough about how to solve his problem that he always asks relevant questions.

The second assumption limits the questions IS can ask to those which the system can answer. But even an expert system has limited domain knowledge. Furthermore, in a rapidly changing world, knowledgeable users may have more accurate information about some aspects of the domain than does the system. For example, a student advisement system may not be altered immediately upon changing the teacher of a course. A cooperative system should recognize its limited knowledge and reason with it to provide whatever pertinent, helpful information it can.

The third assumption restricts IS to utterances that are clear, precise, and accurate. For example, it eliminates the possibility that IS might say he is a junior, when in fact he is three credits short of junior standing, thereby leading the system to erroneously infer that IS is eligible for certain programs or awards. But human information-seekers are often imprecise, especially when they are not aware that small perturbations in the data can be significant.

The fourth assumption says that the system never makes an error in inferring IS's plan. But even in the simplest cases, the system must hypothesize how individual utterances relate to one another. Such decisions select from among multiple possibilities and are a potential source of error.

Pollack (1987) argues against plan inference systems making the first two assumptions, because they prevent

the system from inferring plans which the user believes he can pursue but which are novel (to the system) or invalid. However, there is another implication of relaxing the appropriate query assumption that is not considered by Pollack: IS may ask an irrelevant question that seems perfectly reasonable to the system, thereby leading the system to develop incorrect beliefs about IS's objectives. Consider, for example, a student advisement system. If only B.A. degrees have a foreign language requirement, the query

> "What courses must I take to satisfy the foreign language requirement in French?"

may lead the system to infer that IS is pursuing a bachelor of arts degree. If only B.S. degrees require a senior project, then a subsequent query such as

> "How many credits of senior project are required?"

is problematic. Either the second query is inappropriate to IS's goal of obtaining a bachelor of arts degree (Pollack 1986), or the system's context model does not accurately reflect what IS wants to do. Note that, in either case, the user has a misconception; but in the latter case, the misconception went undetected and was allowed to introduce errors into the system's context model.

Traditional natural language plan inference systems also make the third and fourth assumptions, which, together with the first two, guarantee that the underlying plan inferred by the system and the task-related plan under construction by IS are never at variance with one another. If we want systems capable of understanding and appropriately responding to naturally occurring dialog, natural language interfaces must be able to deal with situations where those assumptions are not true.

Grosz (1981) claimed that miscommunication can occur if both dialog participants are not focused on the same subset of knowledge. Joshi (1982) contended that successful communication requires that the mutual beliefs of the dialog participants be consistent. Extending this to inferred plans, we claim that a successful cooperative dialog requires that the system's beliefs about IS's plan be consistent with what IS is actually considering doing. But clearly it is unrealistic to expect that the system's model will always be correct, given the different knowledge bases of the two participants and the imperfections of communication via dialog.

Thus we need a repair mechanism that attempts to detect inconsistencies in the models and repair them whenever possible. This view is supported by the work of Pollack, Hirschberg, and Webber (1982). They suggested that expert-novice dialogs could be viewed as a negotiation process, during which not only an acceptable solution is negotiated, but also understanding of the terminology and the beliefs of the participants. The context model is one component of the system's beliefs, as is its belief that this model accurately reflects the plan under construction by IS.

5.1 RELATED WORK

Several research efforts have addressed problems related to plan disparity. Kaplan (1982) and McCoy (1986) investigated misconceptions about domain knowledge and proposed responses intended to remove the misconceptions. However, such misconceptions may not be exhibited when they first influence the information-seeker's plan construction; in such cases, disparate plans may result, and correction will entail both a response correcting the misconception and further processing to bring the system's context model and the plan under construction by the information-seeker back into alignment.

Allen's plan inference system (Allen et al. 1980) could accommodate some user misconceptions. It did not expressly eliminate invalid plans but instead weighted them less favorably than valid ones. However, his model did not consider how potential user misconceptions might affect the partial plan inferred by the system.

Pollack (1986) studied removal of the appropriate query assumption of previous planning systems. She proposed a richer model of planning that regarded plans as mental phenomena and explicitly reasoned about the information-seeker's possible beliefs and intentions. She addressed the problem of queries that indicated the information-seeker's plan was inappropriate to his overall goal, and attempted to isolate the erroneous beliefs that led to the inappropriate query. However, queries deemed inappropriate by the system may signal phenomena other than that the query is inappropriate to what the user really wants to do. For example, the information-seeker may have shifted focus to another aspect of the overall task without successfully conveying this to the system, the system's context model may have been in error prior to the query, or, as noted by Pollack (1987), the information-seeker may be addressing aspects of the task outside the system's limited knowledge.

Pollack was concerned with issues that arise when the information-seeker's plan is incorrect due to a misconception exhibited by the current query. She assumed that, immediately prior to the user making the problematic query, the system's partial model of the user's plan was correct. We argue that since the system's inference mechanisms are not infallible and communication itself is imperfect, the system must contend with the possibility that its inferred model does not accurately reflect the user's plan. Previous research has failed to address this problem.

Schmidt, Sridharan, and Goodson (1978) proposed a hypothesize-and-revise paradigm for inferring a user's goal by observing his non-communicative actions. They formulated a set of revision critics for altering a plan upon observing actions that conflict with expectations, but failed to provide any principles or mechanism for selecting the appropriate revision. Although the model

of plan recognition for an office environment formulated by Carver, Lesser, and McCue (1984) attempted to repair the inferred plan when actions inconsistent with it were observed, it did not reason about where the plan might be wrong, but merely backtracked to select another interpretation.

5.2 AN APPROACH TO ROBUST PLAN RECOGNITION

Our analysis of naturally occurring dialog suggests that a plan recognition framework for handling disparate plans should include four phases:

1. Detecting clues to possible disparity between the system's context model and the user's actual goals and plans for accomplishing them. For example, expressions of surprise at the system's response and what appear to be major unsignaled shifts in focus of attention should lead the system to suspect that its context model might be in error.
2. Reasoning on the system's context model and the system's domain knowledge to hypothesize the source of these disparities.
3. Negotiating with the user to isolate the errors. The negotiation phase should be guided by the system's hypothesis about the source of errors in the context model.
4. Appropriately repairing the context model, as indicated by the negotiation dialog.

We believe that the knowledge acquired from the dialog and how it was used to construct the context model are important factors in hypothesizing the cause of disparity between the system's context model and the actual plan under construction by the information-seeker. Natural language systems must employ various techniques such as focusing heuristics and default rules for understanding and relating dialog in order to do the kind of inferencing exhibited in dialogue transcripts and provide the most helpful responses. But confidence in individual components of the resultant context model appears to be important in hypothesizing errors. We contend that the system's context model should be enriched, so that its representation of the plan inferred for the user differentiates among its components according to the support that the system accords each component as a correct and intended part of that plan. The system can then reason on this enriched context model to hypothesize the most likely sources of suspected disparities.

For example, if the system believes that the information-seeker intends the system to recognize from his utterance that G is a component of his plan, then the system can confidently add G to its context model. Components that the system adds to the context model because of the system's domain knowledge should be less strongly believed. This distinction resembles *intended recognition* versus *keyhole recognition* (Cohen et al. 1981). Intended recognition is the inference of those goals and plans that an agent *intends* to convey. Keyhole recognition is the inference of an agent's goals

and plans by unobtrusively observing the agent, as if through a keyhole. Intended recognition is essential in communicative situations (Cohen et al. 1981), since the listener must identify the intended meaning of a speaker's utterance.

Our analysis of naturally occurring dialog suggests keyhole recognition is often critical to expand beliefs about what the information-seeker is trying to do and how it should be done. For example, if CS180 is an introductory course restricted to majors in computer science and electrical engineering, then the system might infer from the utterance

"Can you tell me what time CS180 meets?"

not only that the user wants to know the meeting time for CS180, but also that the user is a computer science or electrical engineering major. However, the user may intend the system to recognize the first goal, but it is questionable whether the user actually intends the system to recognize that the user is pursuing a major in computer science or electrical engineering. This latter inference is based on the system's beliefs about who can take CS180—knowledge that the user may not have. Therefore, since the user may not have intended to communicate these components, they are more likely sources of error than components that the user intended the system to recognize.

The particular rules used to add a component to the context model should affect the system's faith in that component as part of the information-seeker's overall plan. For example, since default inference rules and focusing heuristics select from among multiple possibilities, they add components that are likely sources of suspected errors.

We believe that if a plan recognition system builds such an enriched context model, uses it to hypothesize the source of suspected errors in the model, and attempts to negotiate with the user to isolate and repair its model, the system will be able to handle a much larger set of dialogs than can current models of plan inference, and will be likely to produce responses resembling those found in transcripts of naturally occurring information-seeking dialogs.

6   CONCLUSIONS AND CURRENT RESEARCH

A cooperative natural language system must attempt to infer the underlying task-related plan motivating the information-seeker's queries and use this plan to provide cooperative, helpful responses. The system's model of this plan, which we call a *context model*, is one component of a user model. We have presented a strategy for dynamically inferring the context model from an ongoing dialog, and have shown how this model can be used to handle one class of problematic utterances—the set of utterances that violate the pragmatic rules of the system's world model. Our strategy, motivated by Grice's theory of meaning and maxim of

relevance, often enables the system to deduce the information-seeker's intended meaning, thereby allowing the dialog to continue without interruption.

However, the assumptions underlying current plan inference systems are unrealistic and must be removed. We contend that a natural language system must be able to detect and recover from discrepancies between the system's context model and the actual plan under construction by the user, and have suggested that handling disparate plans requires an enriched context model that differentiates among its components according to the support it accords each component as a correct and intended part of the information-seeker's plan.

7 ACKNOWLEDGMENTS

REFERENCES

Allen, James F. and Perrault, C. Raymond 1980 Analyzing Intention in Utterances. *Artificial Intelligence* 15: 143–178.

Carberry, Sandra 1985 A Pragmatics Based Approach to Understanding Intersentential Ellipsis. In *Proceedings of the 23rd Annual Meeting of the Association for Computation Linguistics*, Chicago, IL: 188–197.

Carberry, Sandra 1986 TRACK: Toward a Robust Natural Language Interface. In *Proceedings of the Sixth Canadian Conference on Artificial Intelligence*, Montreal, Quebec, Canada: 84–88.

Carberry, Sandra 1986 User Models: The Problem of Disparity. In *Proceedings of the 11th International Conference on Computational Linguistics*, Bonn, West Germany: 29–34.

Carver, Norman F.; Lesser, Victor R.; and McCue, Daniel L. 1984 Focusing in Plan Recognition. *Proceedings of the Fourth National Conference on Artificial Intelligence*, Austin, Texas: 42–48.

Chang, C.L. 1978 *Finding Missing Joins for Incomplete Queries in Relational Databases*. Technical Report RJ2145, IBM Research Laboratory, Yorktown Heights, NY.

Cohen, Philip R.; Perrault, C. Raymond; and Allen, James F. 1981 Beyond Question Answering. In W. Lehnert and M. Ringle, (eds.), *Strategies for Natural Language Processing*, Lawrence Erlbaum Associates; 245–275.

Fikes, R.E. and Nilsson, N.J. 1971 STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. *Artificial Intelligence* 2: 189–208.

Grice, H. Paul. 1975 Meaning. *Philosophical Review* 56: 377–388.

Grice, H. Paul. 1969 Utterer's Meaning and Intentions. *Philosophical Review* 68: 147–177.

Grice, H. Paul. 1975 Logic and Conversation. In P. Cole and J.L. Morgan (eds.), *Syntax and Semantics III: Speech Acts*, Academic Press, New York, NY: 41–58.

Grosz, Barbara J. 1977 The Representation and Use of Focus in a System for Understanding Dialogs. In *Proceedings of the International Joint Conference on Artificial Intelligence*, Cambridge, Massachusetts: 67–76.

Grosz, Barbara J. 1981 Focusing and Description in Natural Language Dialogues. In Webber B.; Joshi A.; and I. Sag (eds.), *Elements of Discourse Understanding*, Cambridge University Press, Cambridge, England: 85–105.

Joshi, Aravind K. 1982 Mutual Beliefs in Question-Answer Systems. In N. Smith (ed.), *Mutual Beliefs*, Academic Press, New York, NY: 181–197.

Kaplan, S.J. 1982 Cooperative Responses from a Portable Natural Language Query System. *Artificial Intelligence* 19(2): 165–187.

Litman, Diane 1986 Linguistic Coherence: A Plan-Based Alternative. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York, NY: 215–223.

Mays, E. 1980 Failures in Natural Language Systems: Applications to Data-Base Query Systems. In *Proceedings of the First National Conference on Artificial Intelligence*, Stanford, CA: 327–330.

McCoy, Kathleen F. 1986 The ROMPER System: Responding to Object-Related Misconceptions Using Perspective. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York, NY: 97–105.

McKeown, Kathleen R. 1985 *Text Generation*. Cambridge University Press, Cambridge, England.

Perrault, R. and Allen, J. 1980 A Plan-Based Analysis of Indirect Speech Acts. *American Journal of Computational Linguistics*: 167–182.

Pollack, Martha 1986 A Model of Plan Inference that Distinguishes Between the Beliefs of Actors and Observers. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York, NY: 207–214.

Pollack, M. 1987 Some Requirements for a Model of the Plan-Inference Process in Conversation. In Ronan Reilly (ed.), *Communication Failure in Dialogue*, North Holland: 245–256.

Pollack, M.; Hirschberg, J.; and Webber, B. 1982 User Participation in the Reasoning Processes of Expert Systems. In *Proceedings of the Second National Conference on Artificial Intelligence*, AAAI, Pittsburgh, PA: 358–361.

Reiter, R. 1978 On Closed World Data Bases. In Gallaire H. and Minker J. (eds.), *Logic and Data Bases*, Plenum Press, New York, NY: 55–76.

Rich, E. 1979 User Modeling via Stereotypes. *Cognitive Science* 3(4): 329–354.

Robinson, Ann 1981 Determining Verb Phrase Referents in Dialogs. *American Journal of Computational Linguistics*: 1–18.

Robinson, Ann E.; Appelt, Douglas E.; Grosz, Barbara, J.; Hendrix, Gary G.; and Robinson, Jane J. 1980 *Interpreting Natural Language Utterances in Dialogs about Tasks*. Technical Report TR210, SRI International, Menlo Park, CA.

Schmidt, C.F.; Sridharan, N.S.; and Goodson, J.L. 1978 The Plan Recognition Problem: An Intersection of Psychology and Artificial Intelligence. *Artificial Intelligence* 11: 45–82.

Sidner, Candace L. 1981 Focusing for Interpretation of Pronouns. *American Journal of Computational Linguistics*: 217–231.

Sidner, Candace L. 1983 What the Speaker Means: The Recognition of Speakers' Plans in Discourse. *Computers and Mathematics With Applications* 9(1): 71–82.

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

Sowa, J.F. 1976 Conceptual Graphs for a Database Interface. *IBM Journal of Research and Development*: 336–357.

Thompson, Bozena H. 1980 Linguistic Analysis of Natural Language Communication with Computers. In *Proceedings of the 8th International Conference on Computational Linguistics*: 190–201.

Wilensky, Robert 1983 *Planning and Understanding*, Addison Wesley, Reading, MA.

Wilkins, D.E. 1984 Domain-Independent Planning: Representation and Plan Generation. *Artificial Intelligence* 22: 269–301.

NOTE

1. Allen's and Pollack's models of plan inference are exceptions, as described later.