

TAKING THE INITIATIVE  
IN PROBLEM-SOLVING DISCOURSE

Kevin Matthews  
Kathleen McKeown

CUCS-114-84

Department of Computer Science  
Columbia University  
New York, New York 10027

# TAKING THE INITIATIVE IN PROBLEM-SOLVING DISCOURSE<sup>1</sup>

Kevin Matthews  
Kathleen McKeown

Department of Computer Science  
Columbia University  
New York, New York 10027

## Abstract

Human participants in a conversation display a flexibility in shifting initiative not currently found in natural language processing systems. They frequently take charge of the dialogue to provide necessary, but unrequested information. Based on an analysis of transcripts of student advising dialogues, we present a theory of role shifting that details *the information needed* to take the initiative, *motivations* for doing so, and *expectations* arising from the current discourse state. Motivations for taking the initiative may arise either from the domain task or from the communication requirements of the dialogue itself. To maintain conversational coherence while taking the initiative a system must consider the expectations arising from the current topic and role states. We present data to support our conclusions which have implications for natural language computer systems.

---

<sup>1</sup>This work was partially supported by ONR grant N00014-82-k-0256

## Introduction

In order for a natural language system to engage in a cooperative problem-solving dialogue with a user, the system must not only be able to answer questions, it must also accomplish its own agenda of goals for the session. For instance, an advisory system may need to provide information which the user may not request. In addition, such a system should have the ability to assist the user in specifying a problem, such as by clarifying a question. To achieve these goals, a system must be capable of initiating discourse goals of its own, as well as responding to explicit initiatives of the user. Such a conversational, mixed initiative system differs fundamentally from the more familiar question-answering systems that only reply to questions by retrieving or computing an answer (e.g., [Lehnert 77], [Kaplan 79]).

Previous work on mixed initiative dialogue (e.g., [Allen 83], [Carbonell 79], [Pazzani 83], [Reichman 81]) recognizes the need to guide a participant's next conversational move, but few systems are capable of shifting the direction of the conversation in a unified way to resolve the problem being solved.<sup>2</sup> While Reichman's [81] theory, based on the *context space*, a unit of focus in conversation, does provide a mechanism for managing the topic shifts in a dialogue, she does not address the issue of how a conversant can decide among his current options. One of these options is whether to shift the roles of questioner and responder. In this paper, we extend her theory to managing role shifts and show how a single conversant relates topic and role constraint information to the tasks he or she wishes to accomplish in the dialogue.

## A Sample Dialogue

The present work is based on a study of transcripts of faculty members advising students about course selection. The goal of this study was to formulate a theory of initiative-taking sufficiently precise to provide a computational basis for a conversational computer program. The following transcript illustrates several of the issues involved.

1. s: *Are both of those courses going to be offered next semester?*

---

<sup>2</sup>Note that Pazzani's system, KNOBS, is capable of taking the initiative, but only to obtain missing information that an underlying script dictates is necessary to answer a user's question.

2. p: *that... finite math is going to be offered... and fundamental algorithm is going to be offered, too.*
3. s: *ok but not computability.*
4. p: *and computability? see what number that is. here is by the way your record. 3261? yes that's offered too. so all three classes are offered next semester. except I would suggest that at least fundamental algorithms you take this semester because that's a prerequisite, at least it's very good to know it for many other classes, for many electives.*
5. s: *mmhm*
6. p: *so I would suggest you take at least fundamental algorithm, algorithms this semester.*
7. s: *I know that computability, well both of those courses*
8. p: *mmhm*
9. s: *are not required prerequisites but umm they're recommended. Is it going to be harder for me if I... I mean am I going to be much behind if I haven't taken either of those courses.*
10. p: *mmhm, not really, no. You have to take them; you know that.*
11. s: *I know.*
12. p: *it doesn't really matter what, which semester you take it. What are the electives you want to take?*
13. s: *Ok, I want to take OS.*

### Example 1

This sample provides several examples of the phenomena for which we wish to account. In lines #7-9 the student, s, is in the role of questioner directing the dialogue to accomplish his goal of determining whether it is important to take a certain course at a specific time. At this point the advisor, p, functions in the role of responder as indicated by his passive responses, "mmhm". In line #12, the advisor after rephrasing an answer to an earlier question goes on to introduce the new topic of electives, changing his role from responder to questioner. The new topic of electives proceeds logically from the previous topic of required courses, indicating an underlying organization of domain knowledge that is expressed in the structure of the dialogue. We wish to consider how the advisor arrives at this decision to change topic and role.

## Theory of Context Spaces

Briefly, Reichman's theory of context spaces partitions a conversation into a hierarchy of context spaces, each of which has in focus a topic that may be either of two types, *event* or *issue*. At any moment only the current context space may

be active. Only a small number of relationships is possible between successive context spaces such as: *total shift*, in which the topic changes completely, as in an interruption, *illustrative* in which an event space following an issue space provides an example of the issue in question, or *generating*, in which an issue space summarizes an event space which it follows. Openings and closings of context spaces may be signalled by various linguistic means. These include the use of cue words (such as "ok", "well", "so", and "to return to"), tense shifts, deictic expressions ("this", "that"), pronominalization, and word repetition. In addition, Reichman has described a set of constraints, which she calls "semantics", that limit the ways in which one context space may follow another. For example, an event space which exemplifies the issue of a previous context space may not then be followed by a new issue space generalizing on the event space.

Context spaces and questioner/responder roles are independent of each other. As Reichman shows, more than one context space may occur in a single turn-taking. Moreover, role shifts may occur in a single context space.

### **Extension to Reichman's Theory**

Reichman's theory does not describe what information a single participant will need at each point and how he will use that information to generate an utterance. We describe this information according to the state of the dialogue and the intentions of the speaker. A conversant decides what to say according to the relationships among certain, orthogonal dimensions of the dialogue. These are topic, roles of questioner and responder, and discourse goals which are of roughly two kinds

- **TOPIC** - two basic options are available to the conversant here

1. shift - types of topic shifts include illustrative, generalizing, total, and are governed by Reichman's theory
2. maintain - preserve the same topic in focus

- **ROLES**

1. *questioner*- the conversant who is directing the conversation, usually by asking questions, is the questioner. The options, vis-a-vis the questioner role at any point are:

- a. continue - by exposition or by asking a question.
- b. relinquish - as, by linguistic cue such as a long pause.
- c. acquiesce to seizure by responder - thereby accepting responder role. #13 in example 1 is an example of this. The "ok" of this response signifies only this acquiescence.

2. *responder* - the conversant who is answering questions or permitting the questioner to continue his exposition. Conversant's options in this role are:

- a. continue - answer question or make non-committal response, as in line #5.
- b. seize the speaker's role - as, by asking a question. Both options may be exercised in a single turn as in line #12 of example 1.

- **DISCOURSE GOALS** - at any point, a conversant's purpose may derive from two types of goals:

1. *domain-dependent session goals* - Central to the conversant's participation in this kind of problem-solving dialogue is a flexible set of domain goals to be accomplished in a session ("task goals" [Allen 83]). These goals exist prior to the conversation itself but can be augmented or modified during the session. A goal may permit or require refinement to a more detailed subgoal before being actualized as a discourse goal. In example 1, the student's session goal is to select his courses for the coming semester. This goal could be refined to include a subgoal of verifying that certain prerequisite courses have been completed. One of the advisor's goals is to determine which required courses the student has completed.
2. *dialogue-specific goals* - In addition to the domain-related problems to be solved, other more short-term goals will arise from the dialogue itself ("communication goals" [Allen 83]). For instance, a conversant may simply respond directly to a question, seek additional information before answering it, or decide to clarify a question that is obscure. Additional problems of this type are described in Webber [84].

Both domain-dependent session goals and dialogue-specific goals are dynamically ordered by the conversant according to urgency. Such goals may be

satisfied explicitly or in the process of satisfying another goal. In the former case, the conversant will need to plan and execute a sequence of one or more speech acts to satisfy the goal. Such a speech act may range from a single non-word, as "mmhm", to a series of interacting questions. In this way, either a session goal or dialogue-specific goal leads to a discourse goal, a planning unit for a conversant which is not to be confused with a context space which is a unit of focus for both participants. A planned discourse goal may be interrupted by either conversant without necessarily being abandoned, requiring stacking of discourse goals. Thus, a discourse goal may be in any of the following states:

- a. initiated
- b. continuing
- c. completed
- d. abandoned
- e. suspended
- f. returned to

Of course, not every possible discourse goal admits the possibility of each state. The "mmhm" response is completed in a single utterance.

### **How State Information May Be Used**

A conversant's choices at any point depend upon the state of the conversation, which is to say, the state of the topic and roles, and his intentions or discourse goals. The state of the topic and roles combine to place a mutually recognizable expectation on the turn-taker. The conversant may disregard what is expected of him at any point in the dialogue and choose to pursue a discourse goal he feels is more important, but if he does so he uses his knowledge of the prevailing expectations to choose the appropriate linguistic means to introduce his discourse goal. For example, if the student's last utterance was one of a sequence of questions about the content of a certain elective course, then there is a relatively high expectation that the advisor will continue as responder on the same subject. Should the advisor decide it is more important to determine whether the student has satisfied the prerequisite for the course, he has the option of responding first to the pending question and then seizing the questioner role for his goal or seizing it immediately without responding. Seizing the questioner role without first responding is so contrary to expectations that it requires a more forceful linguistic means to

accomplish it, such as an explanatory remark.

The state of the conversation, the topic and roles, may on the other hand influence the conversant's choice of which discourse goal to pursue. A discourse goal that is semantically close to the current topic may be chosen in preference to another goal which would otherwise take precedence. Initiating a new discourse goal when the conversant is already in possession of the speaker's role likewise requires less disruption of prevailing expectations and therefore less linguistic manipulation to accomplish.

Given this information, we propose several motivations for seizing the questioner role and possibly changing topic. We note three reasons for shifting role, showing how the linguistic means used depend on the state of the dialogue, and cite examples supporting these reasons from the student-advisor transcripts.

1. Need to clarify question before answering it, as in example 2.

*s: I only took Pascal. If I take that would I get a credit  
for it? It's a beginning course.*  
*p: To take another 1000 [level course]?*  
*s: uh huh.*  
*p: no.*

#### Example 2

In this case the topic remains the same although the role shifts. Moreover, there is a high expectation that the questioner, p in this case, will respond to the original question eventually. A number of systems are capable of this type of clarification, most notably the RENDEZVOUS system [Codd 78], without being aware of the distinction in speaker roles. Establishing the questioner/responder distinction provides a basis for further extending the occasions for initiative-taking for other motivations. Note that in this case, the shift is not an abrupt one as topic does not change, and thus the shift in roles does not require the use of strong linguistic cues.

2. the questioner's immediate discourse goal having been satisfied, the responder may introduce a domain goal that proceeds logically from the recent topic and which he feels is necessary for solving the questioner's underlying domain goal. This occurs in line #12 of example 1. Here, the questioner's immediate goal is to determine whether he will be behind if he doesn't take certain required courses next semester. His

underlying goal, however, is to decide what to take next semester. The introduction of electives by the advisor (p) is related to the topic of required courses since electives are a logical postcondition of required courses. Furthermore, discussion of electives is required at some point to satisfy the student's underlying goal. Note that since the new topic is related to the old and the student's immediate discourse goal has been satisfied, the shift is not an abrupt one and no forceful linguistic cues are used to accomplish it.

3. the responder's goal urgently conflicts with the questioner's as in example 3 below.

*s: ... how about computer arts-video games. Which one do you think will be better?*

*p: Well it depends on what you are interested in. Computer art-video games actually it's an introduction to computer graphics.*

*s: mmhm.*

*p: And if you are interested in computer graphics this is your chance to learn about it. Uh, if you are interested in computer networks you should take the other one. Let me first check up. Do you have all your prerequisites? ...*

### Example 3

Here the advisor seizes the questioner role to shift the discussion from high level elective courses to lower level required courses. The advisor's goal is logically prior to the student's and has been raised in urgency by the preceding discussion. An abrupt shift of this sort does not carry the same expectation (as 1 above) that the conversation will return to the interrupted topic. Since the new topic is not related to the old, the shift violates the expectations of the other participant and the advisor must use stronger linguistic cues to signal the shift of roles ("Let me first check up")

So we see that, for motivations arising both from domain and communications problems, a cooperative system needs to have the full capability to shift both its role and the current topic and to introduce the shift using linguistic cues based on prevailing coherence constraints.

## **Future Work**

We have noted some motivations for shifting roles in a cooperative problem solving conversation and plan on investigating others through additional examination

of the transcripts. We have shown that expectations depend to some extent on the relationships between topic and role. We intend to investigate and order these expectations in relation to other expectations such as those that derive from surface structure (e.g., whether a direct or indirect question was used). These theories will be tested by an implementation in the form of a conversational program that helps solve problems in the student advising domain. Implementation of this system is in the preliminary stages.

Reichman's work may be extended in another direction by examining the semantic components of the concept of topic. While Reichman defines topic as a single value for a frame slot, we expect that a topic consists of several dimensions such as point of view and level of abstraction and that these dimensions may influence topic shift as well.

### Summary

Given that a system must be able to take the initiative in a cooperative dialogue, we have shown why and when such a decision may be made. The motivation for initiative taking comes from two sources: domain-dependent session goals and communication goals arising from the dialogue. Making the decision to take the initiative requires consideration of discourse topic and the state of the speaker/responder roles. These two dimensions interact with each other, and possibly with other features of the conversation, to create expectations governing possibilities for shifting roles. Knowledge of these expectations allows a speaker to select the appropriate linguistic means to preserve conversational coherence. By indentifying and defining these motivations, expectations, and components of discourse state information we provide the means for incorporating initiative-taking in a natural language system.

### References

- Allen, J.F., "Argot: A System Overview", *Comp. & Maths. with Appl.* Vol. 9, No. 1, pp. 97-109, 1983
- Carberry, S., "Tracking User Goals in An Information Seeking Environment", *Proc. AAAI*, 1983
- Carbonell, J.G., "Subjective Understanding: Computer Models of Belief

Systems", UMI Research Press, Ann Arbor, Michigan 1981

Codd, E. F., et al., "Rendezvous Version 1: An experimental English-Language Query Formulation System for Casual Users of Relational Databases, IBM Research Laboratory, San Jose, Ca., Tech. Report RJ2144(29407), 1978

Kaplan, S. J., "Cooperative Responses from a Portable Natural Language Database Query System.", Ph.D. Dissertation, U. of Penn., Phila, Pa., 1979

Lehnert, W.G., *The Process of Question Answering*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1977

Pazzani, M.J., "Interactive Script Instantiation", *Proc. AAAI*, 1983

Reichman, R., "Plain Speaking: A Theory and Grammar of Spontaneous Discourse", Report No. 4681, Bolt Beranek and Newman Inc. June 1981

Reichman, R., "Conversational Coherency", Tech. Report #95., Harvard University and Bolt Beranek and Newman Inc., 1978

Schank, R.C. and Wendy Lehnert, "The Conceptual Content of Conversation", Research Report #160, Yale University, 1979

Webber, B.L. and Tim Finin, "In Response: Next Steps in Natural Language Interaction", to appear in W. Reitman (ed), *Artificial Intelligence Applications for Business*, Norwood NJ Ablex Publ. Co., 1984