Proc. 13th Florida Artificial Intelligence Research Symp. (FLAIRS-2000) Orlando, pp. 60–64.

Turn Planning in CIRCSIM-Tutor

Feng-Jen Yang, Jung Hee Kim, Michael Glass, Martha Evens

Computer Science Department
Illinois Institute of Technology
Chicago, IL 60616
{fengjen, janice, glass}@steve.iit.edu, evens@iit.edu

Abstract*

Turn planning is not well-studied in natural language dialogue systems. The current version of CIRCSIM-Tutor attempts to imitate human tutorial dialogue. It generates comprehensible conversation but sounds unnatural. We propose to add a turn planner to improve intra-turn coherence by overall planning of variable references, discourse markers, acknowledgments and softeners. This paper specifies the architecture of the turn planner, which we believe will improve the rhetorical structure in the CIRCSIM-Tutor domain to a reasonable level. Since the idea is not domain dependent, it is potentially applicable to most natural language dialogue systems, especially Socratic dialogue systems.

Introduction

CIRCSIM-Tutor is a natural language-based intelligent tutoring system in the domain of reflex control of blood pressure. Tutoring sessions are organized around solving problems. In the course of solving a problem the student fills in a table with qualitative predictions of changes in several physiological variables, then the tutor starts a Socratic dialogue to remediate the student's errors. CIRCSIM-Tutor is designed to imitate in simplified form both the pedagogical and linguistic behavior of human Socratic tutors.

In this paper we discuss issues of turn planning in CIRCSIM-Tutor: how the various tutor utterances that comprise a single dialogue turn might be combined to produce more fluent dialogue.

CIRCSIM-Tutor plans the tutor's utterances as individual sentences. Each sentence realizes some tutorial or dialogue goal, for example passing judgment on the student's answer, giving a hint, or asking the next question. We have consulted transcripts of human tutors for guidance in the planning process. The human tutoring transcripts are the source of rules that determine the machine's decisions on issues such as what to teach next and how to teach it, how to adjust the tutoring to various student responses, what topics should be elicited from the student and what topics should be ignored, and the language of individual sentences. However the dialogue is generated one sentence at a time. As a result, it sometimes sounds unnatural. Figure 1 contains a fragment of dialogue generated by CIRCSIM-Tutor while in use at Rush Medical School in November, 1998.

T: What are the determinants of Cardiac Output?
S: hr sv
T: Correct, Cardiac Output is determined by Heart Rate and Stroke Volume.
Which determinant is dominant in this case?
S: hr
T: Sorry, the important determinant of Cardiac Output in this case is not Heart Rate.
The important determinant of Cardiac Output in this case is Stroke Volume.
HR didn't change yet.
Is the relationship from Stroke Volume to Cardiac Output direct or is it inverse?
S: direct
T: Good, you got the correct answer.
What is the correct value of Cardiac Output?
S: 0
T: Sorry, the value of Cardiac Output is changed.
Consider the value of Stroke Volume.
What is the correct value of Cardiac Output?
S: down
T: Correct, the value of Cardiac Output is decreased.
What are the determinants of Mean Arterial Pressure?
S: co
T: Right, Cardiac Output is one determinant of Mean Arterial Pressure.
(etc.)

Figure 1. CIRCSIM-Tutor Dialogue Without Turn Planning

Specifically addressing some of CIRCSIM-Tutor's examples of bad textual cohesion, Freedman (1996c) introduced *turn planning* to the CIRCSIM-Tutor project. Freedman (1996b) further described the distinction between two different domains of planning Socratic dialogue: one devoted to satisfying the goals of tutorial dialogue (called the *discourse planner* here), and the turn planner which is devoted to linguistic goals not solely dependent on tutorial structure. The division comes about because tutorial goals and dialogue turns are not congruent:

* This work was supported by the Cognitive Science Program, Office of Naval Research under Grant No. N00014–94–1–0338 to Illinois Institute of Technology. The content does not reflect the position or policy of the government and no official endorsement should be inferred.
Copyright © 2000 American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

one tutorial goal may be realized by several dialogue turns or a fragment of a dialogue turn, conversely a dialogue turn may contain parts or all of several different tutorial goals. Cawsey (1992) effects a similar distinction in the EDGE dialogue-based explanation system. EDGE has "content planning rules" to plan explanations and "dialogue planning rules" which, among other tasks, integrate the explanations into smooth dialogue turns.

The tutor's penultimate turn in Figure 1 (reproduced below) contains a clear example of why tutorial discourse structure does not always match dialogue turn structure. The first sentence results from the end of tutoring one variable while the next results from the beginning of tutoring the next:

A) Correct, the value of Cardiac Output is decreased.
What are the determinants of Mean Arterial Pressure?

That the two sentences serve widely separate tutorial goals can be more clearly expressed like this, using the discourse marker *now then*:

B) Correct, the value of Cardiac Output is decreased.
Now then, what are the determinants of Mean Arterial Pressure?

However in the absence of any other information, normal conversational implicature might incorrectly lead the student to temporarily believe that the two sentences are successive steps in service of the same Socratic goal, which can be expressed with the discourse marker *and*:

C) Correct, the value of Cardiac output is decreased.
And what are the determinants of Mean Arterial Pressure?

One purpose of the turn planner in CIRCSIM-Tutor is to be able to add discourse markers appropriately so that our system generates B) or C) instead of A). Which discourse marker to insert, *and* or *now then*, is determined by the relationships of the two sentences within one turn to the overall tutoring goals. The idea is to generate turns as a coherent whole, not a sentence at a time.

Previous Related Research

Like other interactive dialogue systems, CIRCSIM-Tutor plans text to achieve a discourse goal, and then plans additional text interactively according to the student's response (Woo 1991 and Freedman 1996a). It also formulates the discourse plan in advance and then executes it incrementally. So, during the turn planning, a pre-formulated discourse plan is often being executed partially before it completes, and the turn planner is always prepared to replace the partial plan with a new one or replan according the student's feedback. These features make the turn planning job closely related to both opportunistic planning and incremental planning.

Opportunistic Planning

In the field of opportunistic planning, an opportunity is a possibility subject to a favorable combination of circumstances. This combination of circumstances makes the next step unexpected and unpredictable. An efficient way to plan under such uncertainty is to choose simple plans and adapt them whenever unpredicted circumstances are encountered.

ILEX (the Intelligent Labeling Explorer) is an example of an opportunistic text generation system. It generates a dialogue between a visitor browsing an electronic museum catalog and an agent helping to guide that visitor (Mellish et al. 1998). It retrieves items from the catalog automatically and generates a sequence of descriptions to reflect the interests of the visitor. A feature of its planning is to follow chains of relationships, one fact begetting another, with the ultimate determination of which facts to express being decided by the interests of the visitor and the importance of the fact to the agent's own goal.

Examples D) and E) from (Mellish et al. 1998) each show how ILEX can accrete one fact to another by opportunistically following a link in its knowledge base:

D) Arts and Craft jewelry tends to be elaborate.
For instance, this jewel has floral motifs.
E) This jewel was designed by Jessie M. King.
King worked in London.

If we treat the primitive dialogue acts of CIRCSIM-Tutor as similar to the facts in ILEX's world, the similarity of ILEX to our turn planning is that the choice of the next primitive is always dependent on the combination of all the previous primitives and the current primitive. The principal difference is that the discourse of CIRCSIM-Tutor is organized by schemata while the discourse of ILEX is determined by exploring related facts in its knowledge base.

Incremental Planning

As the name indicates, incremental planning uses planning rules to incrementally expand goals into subgoals. It assumes that achieving one subgoal will not destroy the effects of other subgoals. So, the original goal can be carried out incrementally.

EDGE (the Explanatory Discourse GEnerator) is an example of an incremental text generation system used to generate explanatory dialogue about electronic circuits (Cawsey 92). Since these explanations are interactive, assumptions about the user's background and the current focus may change during the process of the explanation. So, too much detailed planning may be unnecessary and redundant. An important planning idea of the EDGE system is not to commit to the details of the explanation before it has to.

The discourse planning in the EDGE system proceeds incrementally. The primitive action is executed as soon as it is planned. In order to avoid redundancy, the planning expands a goal into subgoals in a depth-first hierarchical manner. When the system is given a topic to explain, it

places this topic on the agenda. As the planning proceeds, it selects a goal from the agenda and executes it, if it corresponds to a primitive action. Otherwise, it selects a planning rule to satisfy this goal and expands some subgoals on the agenda according to this rule. After all subgoals are satisfied in a given order, the original goal is satisfied as well.

The similarity of EDGE to our turn planning is that the discourse plan is carried out in the depth-first hierarchical manner. Once the turn planner has accumulated enough dialogue primitives, it will generate a tutorial turn right away. In this manner, our turn planner is fulfilling a tutorial goal by incrementally satisfying its subgoals.

Putting Together

Planning discourse at the level of a dialogue turn is not a well-studied idea in natural language dialogue systems. The justification for adding this level of planning is to have more fluent and coherent tutorial dialogue.

Reiter's (1994) synthesis of natural language generation architectures bears a similarity to the planners in CIRCSIM-Tutor. He argues that content selection and planning of rhetorical structure occur together. Our discourse planner indeed performs these operations in concert. Reiter argues for a pipelined architecture, and we propose that the turn planner be interposed between our discourse planner and our surface realization step forming a pipeline similar to Reiter's. What novelty is in the turn planner comes from the fact that it is knitting together utterances that are not necessarily related. In the monologue generation systems discussed by Reiter, this situation seems not to occur.

Since the idea is new, the turn planner has to glean methodologies from other levels of discourse planning. Two of the most important methodologies applicable to the CIRCSIM-Tutor domain are opportunistic planning and incremental planning. The turn planner adopts opportunistic strategies to plan the next tutorial turn according to the student's response and uses the incremental strategies both to accumulate the dialogue primitives within a tutorial turn and to carry out the pedagogical goals incrementally.

Turn Planner in the CIRCSIM-Tutor Context

Most of the research in the CIRCSIM-Tutor project is based on the study of transcripts of keyboard-to-keyboard tutoring sessions taught by two physiology professors experienced in tutoring the topic. The students were first-year medical students who were taking the physiology class. We have about fifty one- and two-hour sessions, consisting of more than 5000 turns.

The architecture of CIRCSIM-Tutor described in this section started with Freedman's (1996a) description of tutorial goal structure and discourse structure in the human tutoring transcripts, together with the architecture that she designated to support these structures in a computer tutor. Extending that analysis, Kim and Freedman produced a more comprehensive set of goals that we are using as plan operators (Kim 1999, Kim et al. 1998a).

In CIRCSIM-Tutor the discourse planner provides the input to the turn planner, so we describe its salient features here. Plans are based on schemata (Freedman 1996b). One schema describes a way to satisfy a single goal in terms of sub-goals and alternative goals. The essential features in our application are:

1. The schemata are mostly hierarchical: the subgoals of a schema are satisfied by lower-level schemata.
2. The highest level schema covers solving one problem, most intermediate level schemata cover multiple dialogue turns, the lowest cover one sentence or two or three. A proposed curriculum planner (Cho and Evens 1999) picks the procedure and makes other higher-level decisions.
3. There are two primitive operators, *inform* and *elicit*, which satisfy the lowest level goals by describing a single utterance.
4. The plan changes in response to student input. For example the schema for the *tutors-variable* goal has several alternative sub-schemas for different methods of tutoring that variable. If a student gives an incorrect response to a question, an alternative method schema might be tried or new goals might be inserted to get the student back on track (Zhou and Evens 1999).

The discourse planner constructs enough conversation, executing inform operators, until it executes an elicit operator asking a question. The inform operators do not cause sentences to be immediately emitted, instead they are collected in a buffer until the system emits an elicit operator. After executing elicit the planner pauses to await the student's answer. The buffer of utterances constituting one turn of the tutor's dialogue is the primary input to the turn planner.

Also available to the turn planner is the tutorial goal structure that produced each utterance in the turn buffer. Thus the turn planner has some access to the rhetorical structure of the dialogue it is processing, e.g. whether two adjacent utterances are in service of the same tutorial goal and if so how they are related.

The output of the turn planner is processed by the sentence generator. Based on an analysis of selected sentences in the transcripts (Kim et al. 1998b), the sentences in CIRCSIM-Tutor's repertoire can be described by a small number of features, for example:

<dm> <soft> what is the value of <var> <stage>?

where the place holders are:

<dm> stands for optional discourse marker
<soft> stands for an optional politeness idiom,
e.g. "can you tell me"
<var> stands for a variable name or abbreviation,
or a pronoun
<stage> stands for an optional prepositional
phrase denoting the stage of the problem

A final realization this sentence might be "Now can you tell me what is the value of TPR in RR?"

Many of the features needed for describing a sentence correspond to features that the turn planner needs to determine.

Figure 3 illustrates where turn planning fits in the process of generating CIRCSIM-Tutor dialogue.

Some Phenomena Handled by the Turn Planner

An evaluation of CIRCSIM-Tutor by 50 students at Rush Medical College demonstrated that in the absence of a turn planner its dialogue is already comprehensible. Our goal is to make it more fluent. We therefore have some latitude in picking which discourse features we choose to address. Here are a few that we have studied:

1. Discourse Markers. As we have seen, the introduction of a discourse marker can make clear the logical relationship between two sentences or the relationship of a sentence to the tutorial goal structure. For example, we might attach "so" to the concluding question in the course of tutoring one variable (Yang et al. 2000).
2. Acknowledgments. In human dialogues, acknowledgments of the correctness of student answers are often reduced to a single word, appended to the next sentence, or omitted entirely (Brandle 1998, Spitkovsky and Evens 1993). Whether an acknowledgment can be reduced and attached to the next sentence depends on the syntax of the next sentence and the relation of the next sentence to the answer being acknowledged. However, since the discourse planner emits an acknowledgment separately from the succeeding utterance, it does not assume such reduction is possible. In the dialogue in Figure 1, acknowledgements appear as whole sentences such as "Sorry, the important determinant of Cardiac Output in this case is not Heart Rate." The computer emits the negation of the student's assertion as part of the acknowledgment because the planner does not take into consideration the context in which it occurs.
3. Abbreviated Variable Names. The tutors usually abbreviate variable names. Thus they prefer "CO" to "Cardiac Output." Since they sometimes do not abbreviate, and we believe there may be reasons for occasionally preferring the spelled-out form, the turn planner will have to make this decision (Yang et al. 2000).
4. Softeners. The tutors frequently use various politeness formulas, e.g. asking "Can you tell me, what is …" in place of "What is …."

Broadly, these choices are instances of lexical selection. Since we are using schemata as planning operators, an efficient way of learning the options for lexical selection is by searching for examples of lexical usage in transcripts marked-up with tutoring schemata. We search for instances of the same schema expressed in different ways. After further in-depth analysis of these instances, we can establish rules for lexical selection.

Addressing only the four issues of discourse markers, acknowledgment reduction, variable name abbreviation, and softeners, the dialogue in Figure 1 can be transformed into the passage in Figure 2. Notice that the truth-conditional semantics of the tutor's utterances have not markedly changed, but the dialogue does read quite differently.

T: Do you know what are the determinants of Cardiac Output?
S: hr sv
T: And which determinant is dominant in this case?
S: hr
T: Sorry, the important determinant of CO in this case is SV.
HR didn't change yet.
Is the relationship from SV to CO direct or is it inverse?
S: direct
T: Good.
So what is the correct value of CO?
S: 0
T: No, consider the value of SV.
What is the correct value of CO?
S: down
T: Correct, the value of CO is decreased.
Now, can you tell me what are the determinants of Mean Arterial Pressure?
S: co
T: Right, ….

Figure 2. Dialogue from Figure 1 after Proposed Turn Planning

Conclusion

CIRCSIM-Tutor, which is an attempt to imitate human dialogue, generates comprehensible but stilted conversation. This stiltedness results at least in part from the fact that the discourse planner issues utterances in accordance with a tutorial discourse plan that emits a sentence at a time, without integrating the individual sentences together. We believe that by adding a component that post-processes the sentences within a single dialogue turn we can improve the rhetorical structure of our tutorial turns to a reasonable level. This justification is also compatible with the necessity of paragraph planning (Hovy 1990) and (Mann and Moore 1981).

References

Brandle, S. 1998. Using Joint Actions to Explain Acknowledgments in Tutorial Discourse: Application to Intelligent Tutoring Systems. Ph.D. diss., Computer Science Department, Illinois Institute of Technology, Chicago, IL.

Cawsey, A. 1992. *Explanation and Interaction: The Computer Generation of Explanatory Dialogues*. Cambridge. MA: MIT Press.

Cho, B. and Evens, M. 1999. A Curriculum Planner Model for an Intelligent Tutoring System. In *Proceedings of the Twelvth International Florida AI Research Society Conference (FLAIRS-99)*, Orlando, FL, pp. 197–201.

Freedman, R. 1996a. Interaction of Discourse Planning, Instructional Planning and Dialogue Management in an Interactive Tutoring System. Ph.D. diss., Computer Science Department, Northwestern University, Evanston, IL.

Freedman, R. 1996b. Using a Text Planner to Model the Behavior of Human Tutors in an ITS. In *Proceedings of the Seventh Midwest AI and Cognitive Science Conference (MAICS '96)*, Bloomington, IN. np.

Freedman, R. 1996c, Using Tutoring Patterns to Generate More Cohesive Text. In *Proceedings of the 2nd International Conference on the Learning Sciences*, Evanston, IL, 1996

Hovy, E. 1990. Unresolved Issues in Paragraph Planning. In Dale, R., Mellish, C. and Zock, M., (editors.), *Current Research in Natural Language Generation*. pp. 17–45. Academic Press.

Kim, J.H. 1999. The SGML Markup Manual for CIRCSIM-Tutor. Technical Report, Computer Science Department, Illinois Institute of Technology, Chicago, IL.

Kim, J.H., Freedman, R., and Evens, M. 1998a. Responding to Unexpected Student Utterances in CIRCSIM-Tutor V3.0. In *Proceedings of the Eleventh Florida Artificial Intelligence Research Symposium (FLAIRS '98)*, Sanibel Island, FL, pp. 153–157.

Kim, J.H., Freedman, R., and Evens, M. 1998b. Relationship between Tutorial Goals and Sentence Structure in a Corpus of Tutoring Transcripts. In *Proceedings of the Ninth Midwest AI and Cognitive Science Conference (MAICS '98)*, Dayton, OH, pp. 124–131.

Mann, W. and Moore, J. 1981. Computer Generation of Multiparagraph English Text. *The Journal of Computational Linguistics*, Vol. 7, No. 1, pp. 17–29.

Mellish, C., O'Donnell, M., Oberlander, J. and Knott, A., 1998. An Architecture for Opportunistic Text Generation. In *Proceedings of the 9th International Workshop on Natural Language Generation*, Niagara-on-the-Lake, Ontario, Canada. pp. 18–37.

Reiter, E. 1994. Has a Consensus NL Generation Architecture Appeared, and is it Psycholinguistically Plausible? In *Proceedings of the Seventh International Workshop on Natural Language Generation (INLGW-1994)*, pp. 163–170.

Spitkovsky, J.A. and Evens, M. 1993. Negative Acknowledgments in Natural Language Tutoring Systems. In *Proceedings of the Fourth Midwest AI and Cognitive Science Conference (MAICS '93)*. pp. 41–45.

Woo, C. 1991. Instructional Planning in an Intelligent Tutoring System: Combining Global Lesson Plans with Local Discourse Control, Ph.D. diss., Illinois Institute of Technology, Chicago, IL.

Yang, F., Kim, J., Glass, M. and Evens, M. 2000. Lexical Usage in the Tutoring Schemata of CIRCSIM-Tutor: Analysis of Variable References and Discourse Markers. In Proceedings of Human Interaction with Complex Systems (HICS '00), Urbana, IL, Forthcoming.

Zhou, Y. and Evens, M. 1999. A Practical Student Model in an Intelligent Tutoring System. In *Proceedings of the IEEE International Conference on Tools with Artificial Intelligence*, Chicago, IL, pp. 13–18.

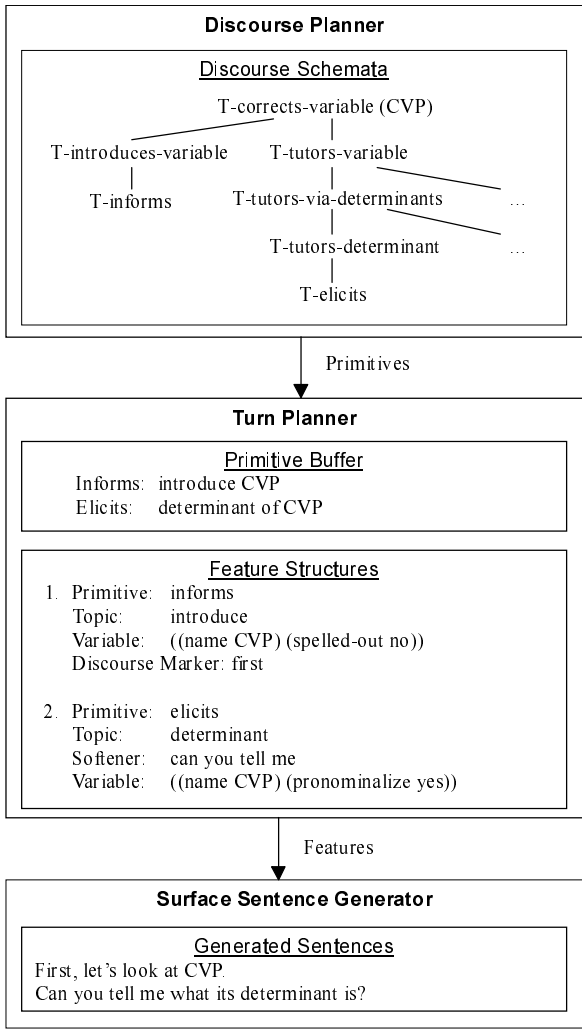

Figure 3: Context of Turn Planner in CIRCSIM-Tutor