# Representing Communicative Action in a Dialogue-Based Intelligent Tutoring System

**Reva Freedman**

Department of CSAM
Illinois Institute of Technology
10 W. 31st Street 236–SB
Chicago, IL 60616
freedman@steve.csam.iit.edu

**Abstract**

We are currently implementing CIRCSIM-Tutor v. 3, a natural-language based intelligent tutoring system (ITS) which tutors medical students on a topic in cardiovascular physiology (the baroreceptor reflex). We outline the text generation process in CIRCSIM-Tutor and illustrate a number of issues we have faced with respect to the representation of communicative action.

## Current Work

CIRCSIM-Tutor is a conversational intelligent tutoring system (ITS) which uses natural language for both input and output. Each tutor turn is constructed as follows:

- Parse the student's input
- Derive an abstract representation of the input
- Plan a tutoring intervention
- Generate a response

CIRCSIM-Tutor contains a global planner for instructional goals, so the tutoring intervention is constructed using parts from the global plan as well as parts which directly respond to the student's previous statement.

Version 3 of CIRCSIM-Tutor, which we are in the process of implementing, differs from many other ITSs in that its underlying architecture uses a text generation paradigm. In other words, the planner's fundamental goal is "generate a conversation resulting in the student knowing <concepts>" rather than "teach <concepts>." Thus the logical forms used for representing both input and output and many of the intermediate forms used by the planner, are representations of communicative action. (The planner is described further in Freedman 1996).

Our basic knowledge representation for instructional planning is a sophisticated form of schema which allows static and dynamic preconditions, recursion and full unification. We have used schemata for this application because we need some control over the way concepts are taught, including decomposition and sequencing.

The following schema is typical (schemata are implemented in Lisp):

```
To teach about any neural variable ?v:
    Teach about mechanism of control of ?v
    Teach about when this mechanism is activated
    Find out if student knows the correct answer yet
```

The schemata were abstracted from transcripts of human tutors with students. The transcripts show the hierarchical structure of instructional planning. Although our formalism in no way requires it, it is interesting to note that regardless of the topic, all of our tree structures contain the same category of information at a given level of the tree. In this respect our analysis shows similar results to that of Kerpedjiev et al. (1997) for a multimedia generation system. The following subtree is typical:

```
Tutor variable which student missed
    Introduce variable to be discussed
    Strategy for tutoring that variable
        Tutor first concept of pattern
            Speech act
        Tutor second concept of pattern
            . . .
```

Each concept can be taught in one of several ways:

- Inform the student about the concept
- Elicit the desired information from the student
- Don't generate text for this instance of this topic

The latter option is used to generate hints to the student by suppressing part of a schema.

If the student answers a question incorrectly, CIRCSIM-Tutor can backtrack at three levels:

- Different linguistic realization of the speech act
- Different speech act
- Different tutoring strategy

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.

## Issues

This summary highlights a number of issues we have identified in trying to represent communicative actions for tutoring. The underlying theme of many of these issues is the interaction during text generation between domain information and linguistic information such as topic, focus and reference.

*What kinds of actions are needed beside sentence-level speech acts?*

Our planner currently includes acts for starting and terminating sections of dialogue. These logic forms, such as the italicized one in the example above, are essential for generating a coherent dialogue. The word *introduce* can be used to represent a pedagogical act or a higher-level discourse act. We are currently using it to refer to a pedagogical act. In the future we would like to support the meaning of *introduce* as a discourse act too. Such an enhancement would include the possibility of directly representing other pragmatic acts, such as "introduce a referent for <object>" or "change the topic to <item>."

At the topic level we can directly represent higher-level actions which the tutor can take, such as *explain*, *define*, and so on. These are implemented using standard primitives such as *inform*. Other actions which the tutor can take, such as giving the student a hint, are best viewed as emergent properties of the dialogue and as such are not directly represented in the planner.

*How should actions be combined into sentences?*

In CIRCSIM-Tutor, the instructional planner determines which speech acts are to be issued and in what sequence. A separate planner, which could be any paragraph planner, is used at the turn level to combine a series of speech acts into a coherent turn. If the generated text contains subordinate clauses, some method, such as rhetorical relations, schemata or prestored trees, must be used to identify what can be combined. For example, in the following sentence, does the subordinate clause have a separate existence in the underlying representation:

> If cardiac output is made to vary, what effect will that have on the central venous compartment?

*How are dialogue acts related to mental states?*

We are currently experimenting with several representations for the tutor's and student's mental states. The following alternatives are listed in order of decreasing assumptions about mental states:

- (student-knows (P x))
- (tutor-believes (student-knows (P x))
- (tutor-has-taught (P x))
- (tutor-has-said (P x))

The difference between basing the knowledge representation on what the student knows and basing it on what the tutor says is closely related to the difference between Chi's verbal analysis method (1997) and the protocol analysis method of Ericsson & Simon (1984).

*What features of the text realization must be supported by the action representation?*

For example, when do we need to represent the difference between closely related expressions such as:

- What is the value of heart rate?
- What happens to heart rate?

Sinclair & Coulthard (1975) point out that there is no unique relationship between a speech act and the form of its realization as text. For example, *elicit* is not restricted to the interrogative:

- What is the value of heart rate?
- Tell me the value of heart rate.
- Now I'd like to know the value of heart rate.

A similar issue arises with respect to indirect methods of expression and politeness formulae, e.g.:

- Could you tell me the value of heart rate?

The representation of discourse markers brings up similar issues:

- So what happens to heart rate?

## References


Chi, M. T. H. 1997. Quantifying Qualitative Analyses of Verbal Data: A Practical Guide. *Journal of the Learning Sciences* 6(3): 271–315.

Ericsson, K. A. and Simon, H. 1984. Protocol Analysis: Verbal Reports as Data. Cambridge, MA: MIT Press.

Freedman, R. 1996. Interaction of Discourse Planning, Instructional Planning and Dialogue Management in an Interactive Tutoring System. Ph.D. diss., Dept. of EECS, Northwestern Univ.

Kerpedjiev, S., Carenini, G., Roth, S., Moore, J. D. 1997. Integrating Planning and Task-based Design for Multimedia Presentation. In International Conference on Intelligent User Interfaces (IUI '97), Orlando, 145–152.

Sinclair, J. M. and Coulthard, R. M. 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils.* London: Oxford University Press.


## Acknowledgments


Many thanks to Jung-Hee Kim, whose meticulous coding of transcripts elucidated some of the patterns described here, and to Professor Martha W. Evens, who has directed the CIRCSIM-Tutor project since its inception.