# Discourse Patterns in Why/AutoTutor

**Eric C. Mathews [1], G. Tanner Jackson[1], Arthur C. Graesser[1], Natalie K. Person[2] and the Tutoring Research Group[1]**

[1] Department of Psychology, University of Memphis, Memphis, TN 38152
emathews@memphis.edu, gtjacksn@memphis.edu, a-graesser@memphis.edu

[2] Department of Psychology, Rhodes College, 2000 N. Parkway, Memphis, TN 38112
person@rhodes.edu

**Abstract**

The interfaces of knowledge management systems will benefit from conversational agents, particularly for users who infrequently use such systems. The design of such agents will presumably share some of the dialog management facilities for systems designed for tutoring. For example, Why/AutoTutor is an automated physics tutor that engages students in conversation by simulating the discourse patterns and pedagogical dialog moves of human tutors. This paper describes how the Why/AutoTutor creates original dialog pathways for the learner. The system chains dialog moves, expressions, and discourse markers to simulate the dialog moves of natural human tutors while still controlling the conversational floor and the learning of the student.

The agents of some knowledge management facilities of the future will be intelligent conversational agents. Conversational agents direct the flow of mixed-initiative dialog in service of mutual goals. These agents prompt the user when to speak and what to say, provide useful feedback, and answer questions. Conversational agents will be particularly useful for infrequent users of a knowledge management system because they need the most guidance in managing interactions with the system.

Animated conversational agents have recently been designed for learning environments and help facilities (Cassell & Thorisson, 1999; Graesser, VanLehn, Rose, Jordon, & Harter, 2001; Johnson, Rickel, & Lester, 2000). These systems have dialog management facilities that hold mixed initiative dialog with the user by generating a variety of content-sensitive discourse moves: questions, answers, assertions, hints, suggestions, feedback, summaries, and so on. The design of these systems presumably have features and components that would directly apply to intelligent agents that control knowledge management systems. The purpose of the present paper is to describe the dialog facilities of one such conversational agent (AutoTutor) which was designed for tutoring. It is an open question as to whether an intelligent tutoring system would need sufficient overlap with the design of a knowledge management system in order to be useful.

## What is Why/AutoTutor?

Why/AutoTutor is the newest in a series of intelligent tutoring systems that engage in a conversation with the learner while simulating the dialog moves of human tutors (Graesser, VanLehn et al., 2001). As in all versions of AutoTutor (Graesser, Person, Harter, & TRG, 2001), the student and AutoTutor collaboratively improve the quality of the student's answers to problems and questions. This is accomplished by participating in a mixed-initiative conversation, distinguishing it from mere information delivery systems.

The latest version of AutoTutor, called Why/AutoTutor, was specifically designed to help college students learn Newtonian conceptual physics. Previous versions of AutoTutor taught college students about computer literacy. Why/AutoTutor is comprised of six major modules: (1) an animated agent, (2) a curriculum script, (3) a speech act classification system (SAC), (4) latent semantic analysis (LSA), (5) a dialog move generator, and (6) a question answering tool (QUEST). Many of these modules of AutoTutor have been discussed rather extensively in previous publications, so they will only be mentioned briefly in this paper (Graesser, VanLehn, et al., 2001; Graesser, Person, et al., 2001; Person, Graesser, Kreuz, Pomeroy, & the TRG, 2001).

Why/AutoTutor's animated agent is a three-dimensional embodied agent that remains on the screen during the entire tutoring session. It was created using Curious Labs' Poser 4 and is controlled using Microsoft Agent software. Dialog moves generated by Why/AutoTutor during the tutoring session are synchronized with the agent's head movements, facial expressions, and hand gestures. The dialog moves serve both conversational and pedagogical functions.

Why/AutoTutor always begins each tutoring session by introducing the functionality of the system, and then begins tutoring the learner using material found within the curriculum script. Our physics curriculum script is a set of organized concepts, misconceptions, good answers, bad answers, teaching points, and question-answer units of various types. The curriculum script serves to organize the topics to be covered during the tutoring session. This content takes on the form of word phrases, sentences, questions, and paragraphs in a structured text format. Presently Why/AutoTutor tutors students on 10 conceptual physics problems. Each of these physics problems is represented in the curriculum script with the following slots of information:

(1) Statement of the problem to be solved, in the form of a question.
(2) A set of expectations in an ideal answer, with each expectation being a sentence in natural language of 10-20 words
(3) A set of tutor dialog moves that express or elicit from the learner each expectation in #2 (i.e., hints, prompts, and assertions)
(4) A set of anticipated bad answers and corrections for those bad answers
(5) A set of physics misconceptions and corrections to those misconceptions
(6) A set of basic noun-like concepts about physics and their functional synonyms in the specific context of the problem.
(7) A summary of the answer or solution
(8) A latent semantic analysis (LSA) vector for each expectation, bad answer, and misconception. LSA is discussed further below.

During a tutoring session Why/AutoTutor provides the student with one of the physics problems and asks the student to provide an answer. Using the information found within the curriculum script, Why/AutoTutor works with the student toward the ideal answer. To aid the student in this process, Why/AutoTutor attempts to extract the desired information from the student by generating prompts and hints for each expectation in the ideal answer for the problem. Throughout this process, Why/AutoTutor uses LSA, to assess the student contributions and decide which portions of the ideal answer are missing. That is, Why/AutoTutor compares material found within the curriculum script to the student's contributions using LSA, a statistical technique that measures the conceptual similarity of two texts (Foltz, 1996; Foltz, Britt, & Perfetti, 1996; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). We have found LSA to be sufficiently accurate in evaluating student contributions, and thus have applied it in Why/AutoTutor (Graesser, Wiemer-Hastings, Wiemer-Hastings, Harter, Person, & TRG, 2000; Olde, Franceschetti, Karnavat, Graesser, & TRG, 2002; Wiemer-Hastings, Wiemer-Hastings, Graesser, & TRG, 1999).

When students interact with Why/AutoTutor they type their contributions into a text entry box on the interface and submit the material by hitting the "Enter" key. Upon submission Why/AutoTutor analyzes the student turn by assessing what type of contribution the student provided. The speech act classifier (SAC) of Why/AutoTutor does this using a parser and a series of rules and transformations that determine into which category the student's contribution falls. The SAC assigns the student's input into one of 23 speech act categories including 17 question categories. Example categories are as follows: Assertion, Yes/No question, Prompt Completion, Causal Antecedent question, and so on. These speech act categories enable AutoTutor to sustain mixed-initiative dialog as well as dictate how Why/AutoTutor will generate pedagogically effective and conversationally appropriate dialog moves. The dialog generation pathways will be discussed more extensively in the third section of this paper.

Once Why/AutoTutor concludes that the student has covered the critical components for a particular problem, the system asks the student if he/she has any further questions and then asks the student to recap what he/she has learned in the problem. This is followed by Why/AutoTutor providing a brief summary of the material. The conversation then proceeds to the next problem in the curriculum script, or the tutoring session concludes.

## Dialog Moves in Tutoring and Why/AutoTutor

Our goal from the onset of the AutoTutor project has been to develop an agent that simulates the dialog moves of human tutors while participating in a natural conversation with the learner. To accomplish this ambitious goal, we examined tutoring strategies used in one-to-one human tutoring (Graesser, Person, & Magliano, 1995; Person, & Graesser, 1999). We developed a set of tutorial dialog moves and rules for Why/AutoTutor that mimics the strategies of human tutors.

Why/AutoTutor uses knowledge construction dialog moves that aid in the development of concepts within the mind of the learner. They include hints, prompts for specific information, elaborative assertions, corrections, and pumps. These dialog moves extract important information about the problem at hand and provide the student with correct information with which to proceed in answering the problem. Why/AutoTutor makes use of five forms of short feedback during tutoring (positive, positive-neutral, neutral, negative-neutral, and negative). These dialog moves inform the learner of the correctness of the contribution provided. Why/AutoTutor makes use of thirteen other dialog moves in the tutoring session to perform specific functions in the conversation, such as summarizing, inviting learner questions, and asking the learner to rephrase answers. In addition to the 23 dialog moves mentioned above, Why/AutoTutor generates several types of discourse markers (e.g., “okay,” “and also”) that serve as transitions to various dialog moves and that demarcate different components of Why/AutoTutor’s response to a student contribution.

## Dialog Move Generation and Conversation Management

Each learner contribution is analyzed in two phases: classification of the speech act into one of 23 categories and the quality of the contribution as measured by LSA. With this information, the system is able to decide which of the dialog pathways it should use. Why/AutoTutor does this in several steps, which are briefly outlined below:

1. Determine student initiative and assess turn on quality
2. Choose appropriate tutor response (feedback, hint, pump, nothing, etc.)
3. Adapt to student initiative and transition
    a. If it is a substantive contribution continue with tutoring
    b. If it is other than a substantive contribution, provide needed information to keep conversation on course
4. Transition back to lesson plan and provide next knowledge construction dialog move

*(Note: This is what generally occurs for most student contribution. However, at special points in the tutoring session very specific paths are taken, such as to end a tutoring session or move on to a new problem)*

For all student contributions Why/AutoTutor uses a set of production rules that are sensitive to the quality of student contributions, as measured by LSA. LSA provides the following parameters for the production rules to use: (a) student assertion quality for the previous student turn, (b) student ability level, induced from the quality of all previous student turns in the entire session, and (c) coverage of the expectations in the curriculum script. The SAC provides the system with information about student intention (i.e. whether the student asked a question or gave an answer). The production rules utilize these parameters to specify the conditions and transitions for various dialog moves. A simplified example of how Why/AutoTutor decides a Dialog pathway is provided in Figure 1.

***Student states: The horizontal and vertical components of the object’s motion are independent of one another***

Why/AutoTutor analyzes the contribution:

- Student is supplying information (Speech Act Classifier Information)
- Quality of student information is good (LSA information)
- Student has not covered topic entirely (LSA information)

Why/AutoTutor selects Dialog Pathway:

- Student information is good quality, so provide positive feedback
- Provide an advancer to maintain conversational floor
- Select type of dialog move (i.e. Hint) to elicit new information not covered
- Insert appropriate transition (i.e. Hint Discourse Marker)
- Provide selected dialog move

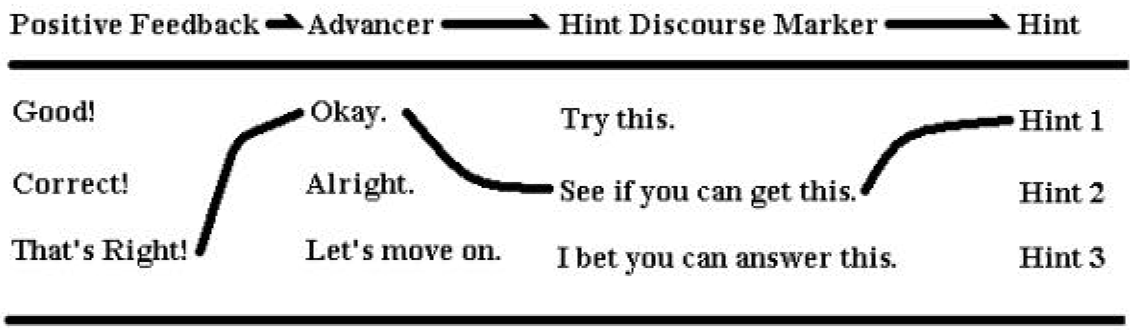
***Why/AutoTutor replies: That’s right! Okay, See if you can get this, what is the vertical force acting on the object in this problem?***

This example, though presented at a superficial level, demonstrates the method used by AutoTutor when selecting dialog pathways. With the information given by LSA and the speech act classifier, the system first selected positive feedback, then maintained the conversational floor using an advancer, then used a hint discourse marker to transition to the hint dialog move, and then generated a hint to elicit more information from the student. This dialog move sequence can then be diagramed as taking the following trajectory: Positive Feedback > Advancer > Hint Discourse Marker > Hint.

This example is actually a simplification. During the tutorial conversation for each tutoring topic, AutoTutor keeps track of which parts of the problem have been covered, using LSA, along with which dialog moves have been previously generated. Why/AutoTutor decides which expectation to cover next after considering what aspect of the ideal answer is missing. There are different algorithms for computing the next expectation to cover, but these are not addressed in the present paper.

From all this information Why/AutoTutor selects its dialog pathway. Pedagogical effectiveness is maintained by proper selection of the dialog moves issued to the student (e.g. feedback, hints, prompts, etc.). Conversational smoothness is maintained by selecting natural transitions between those moves (e.g. discourse markers and advancers).

Figure 1. Example of a Dialog Sequence Chain in Why/AutoTutor.




Note: Hint 1, 2, 3, etc. are different possible questions from the curriculum script.

Linking all of these parts together, Why/AutoTutor generates natural language appropriate for tutorial dialog. The system can be considered a model of dialog chaining. Once the system has all of the assessments, it uses the set of production rules and overarching tutorial dialog goals to select one expectation from the curriculum script to cover and manages the conversation to help the learner to articulate the expectation. Once the system has decided to cover a particular expectation, it produces the sequences of dialog moves and markers in each turn to attempt to elicit the expectation; it may take several conversational turns to do so. So when Why/AutoTutor decides to use a hint dialog move it goes to the curriculum script and selects the actual hint. This contributes to the chaining aspect of the dialog. Figure 1 above provides a simplified version of Why/AutoTutor's chaining for the sequence we used previously: Positive Feedback > Advancer > Hint Discourse Marker > Hint. The computer's response to the student's contribution in the example turn is "That's Right! Okay. See if you can get this: What is the horizontal velocity of the object?" "What is the horizontal velocity of the object?" is Hint 1 from the curriculum script.

In order to understand the generative space of chains provided by Why/AutoTutor, we can compute the number of possible natural language expressions available to the system. For any given dialog move (such as an advancer or discourse marker) there are approximately 30 possible categories of expressions from which to chose and sometimes dozens of natural language expressions per category. There are approximately 4-7 expectations per problem and 25 hints, prompts, and assertions per expectation. Each tutor turn has approximately 2-7 dialog moves. The space of possible sequences of natural language expressions per turn is very large, on the order of millions.

## Evaluation of the Dialog Moves and Pathway Sequences

We recently conducted an evaluation study of Why/AutoTutor on 24 college students. The students interacted with Why/AutoTutor over the course of two 3-hour sessions. The students worked on 10 conceptual physics problems with Why/AutoTutor during this time. In order to insure that our underlying conversational structure was working properly, we examined Why/AutoTutor's dialog moves.

We examined two aspects of the dialogs: (1) the frequency of the dialog move categories used by Why/AutoTutor and (2) the frequency of Why/AutoTutor's major dialog pathways. To simplify our analyses we decided not to include the various categories of discourse markers, because these items do not carry content information about the problem and do not have any obvious pedagogical consequences. Tables 1 and 2 present the results of these two frequency analyses.

## Table 1. Prominent Dialog Moves

| Dialog Move Category | Frequency | % of Total group |
|---|---|---|
| Hint (i.e. Why does the object move with constant velocity?) | 2321 | 14 |
| Prompt (i.e. The object will land on the ____________) | 2095 | 12 |
| Assertion (Velocity is the distance covered per unit time.) | 1861 | 11 |
| Negative Feedback (i.e. No.) | 1325 | 8 |
| Positive Feedback (i.e. Good.) | 3143 | 19 |
| Advancer (i.e. Okay.) | 1594 | 9 |
| Prompt Response Correct (i.e. That's right.) | 1634 | 10 |
| All Other Moves (Except Discourse Markers) | 2943 | 17 |
| **TOTAL** | **16916** | **100** |

**Table 1** above shows the most frequently used dialog moves of Why/AutoTutor. It makes sense that these dialog moves occur with such high frequency in the tutoring sessions since these dialog moves would be most active in the pedagogy of tutoring. Human tutors ask questions, ask learners to fill in content, and provide feedback on answers to further learning.

**Table 2** below shows the most prominent pathways of the approximately 50 measured pathways used by Why/AutoTutor to convey knowledge to the learner. Again, the linkage between prominence and pedagogy in tutorial dialog is manifest in these frequencies. It should be noted that most turns have several dialog moves, not just one dialog move per turn.

## Table 2. Prominent Dialog Move Sequences

| Dialog Move Sequence | Frequency | % of Total |
|---|---|---|
| PositiveFeedback -> PromptResponseCorrect -> Advancer -> Assertion -> Hint | 809 | 14 |
| PositiveFeedback -> Prompt | 783 | 13 |
| NegativeFeedback -> Prompt | 674 | 12 |
| PositiveFeedback -> PromptResponseCorrect -> Assertion -> Hint | 509 | 9 |
| NegativeFeedback -> PromptResponseIncorrect -> Advancer -> Assertion ->Hint | 273 | 5 |
| PositiveFeedback -> Hint | 321 | 5 |
| RephraseAnswer | 238 | 4 |
| PositiveFeedback -> Summarize -> Take_a_break | 217 | 4 |
| NeutralNegativeFeedback -> Prompt | 216 | 4 |
| **Subtotal** | **4040** | **70** |

## Conclusions and Implications

Why/AutoTutor generates many dialog pathways as it adaptively responds to the student during tutoring. It is able to convey information to the learner in a variety of fashions due to its method of chaining expressions, dialog moves, and discourse markers together. These chains serve both pedagogical and conversational goals.

There are two major advantages associated with the model of natural language interaction used by Why/AutoTutor. The first advantage we outlined in section 4: the vast diversity of possible dialog pathways. There are millions of available tutor responses to student contributions, so the likelihood of the system ever repeating itself is close to zero. The second advantage of this model is the fact that it can occur in real time and is computationally feasible. It is possible to theoretically consider a more sophisticated dialog management system that tracks the beliefs, goals, and intentions of the learner, and that dynamically plans utterances in a fashion that caters to these constraints (Rich & Sidner, 1998). However, there are serious computational obstacles of these more sophisticated systems that prevent them from being delivered in practice. The AutoTutor architecture is available until these more sophisticated systems can be developed and scaled up to real world applications. Meanwhile, with these advantages in AutoTutor, we can ensure that Why/AutoTutor is capable of producing diverse, pedagogical and conversationally appropriate tutoring turns that adapt to the student and ultimately result in significant learning gains.

The question remains, however, whether similar systems can be developed in agents for knowledge management. The categories of dialog moves will no doubt be different because the tasks of the users are quite different in knowledge management than in tutoring. However, we do believe it is worth a try to develop an AutoTutor mechanism as a next step.

## Acknowledgements

The Tutoring Research Group (TRG) is an interdisciplinary research team comprised of approximately 35 researchers from psychology, computer science, physics, and education (See http://www.autotutor.org). This research conducted by the authors and the TRG was supported by grants from the National Science Foundation (SBR 9720314 and REC 0106965) and the Department of Defense Multidisciplinary University Research Initiative (MURI) administered by the Office of Naval Research under grant N00014-00-1-0600. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of ONR or NSF.

## References

Cassell, J., and Thorisson, K.R. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. Applied Artificial Intelligence, 13, 519-538.

Foltz, P.W. (1996). Latent semantic analysis for text-based research. Behavior Research Methods, Instruments, and Computers, 28, 197-202.

Foltz, P. W., Britt, M. A., & Perfetti, C. A. (1996). Reasoning from multiple texts: An automatic analysis of readers' situation models. Proceedings of the 18th Annual Conference of the Cognitive Science Society (pp. 110-115). Mahwah, NJ: Erlbaum.

Graesser, A.C., Person, N., Harter, D., & TRG (2001). Teaching tactics and dialog in AutoTutor. International Journal of Artificial Intelligence in Education, 12, 257-279.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialog patterns in naturalistic one-on-one tutoring. Applied Cognitive Psychology, 9, 359-387.

Graesser, A.C., VanLehn, K., Rose, C., Jordan, P., and Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. AI Magazine, 22, 39-51.

Graesser, A.C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & TRG (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. Interactive Learning Environments, 8, 129-148.

Johnson, W. L., & Rickel, J. W., and Lester, J.C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. International Journal of Artificial Intelligence in Education, 11, 47-78

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. Discourse Processes, 25, 259-284.

Olde, B. A., Franceschetti, D.R., Karnavat, Graesser, A. C. & the Tutoring Research Group (2002). The right stuff: Do you need to sanitize your corpus when using latent semantic analysis? Proceedings of the 24th Annual Conference of the Cognitive Science Society (pp. 708-713). Mahwah, NJ: Erlbaum.

Person, N. K., & Graesser, A. C. (1999). Evolution of discourse in cross-age tutoring. In A.M.

O'Donnell and A. King (Eds.), Cognitive perspectives on peer learning (pp. 69-86). Mahwah, NJ: Erlbaum.

Person, N.K., Graesser, A.C., Kreuz, R.J., Pomeroy, V., & TRG (2001). Simulating human tutor dialog moves in AutoTutor. International Journal of Artificial Intelligence in Education. 12, 23-39.

Rich, C., & Sidner, C.L. (1998). COLLAGEN: A collaborative manager for software interface agents. User Modeling and User-adapted Interaction, 8, 315-350.

Wiemer-Hastings, P., Wiemer-Hastings, K., & Graesser, A. C. (1999). Improving an intelligent tutor's comprehension of students with Latent Semantic Analysis. Artificial Intelligence in Education (pp. 535-542). Amsterdam: IOS Press.