Intelligent Tutoring Systems with Conversational Dialogue
Intelligent Tutoring Systems
Overview

- **Intelligent Tutoring System** (ITS) - computer-based training system that incorporate techniques for communicating / transferring knowledge and skills to students.

- ITS = combination of **Computer-Aided Instruction** (CAI) and **Artificial Intelligence** (AI) technology

- Initial research:
  - CAI: 1950
  - ITS: 1970
AUTOTUTOR

• AutoTutor is a web-based intelligent tutoring system developed by an interdisciplinary research team - Tutoring Research Group (TRG);
• This team is currently funded by the Office of Naval Research and the National Science Foundation and it has 35 researchers from psychology, computer science, linguistics, physics, engineering, and education.
• TRG has conducted extensive analyses of human-to-human tutoring, pedagogical strategies, and conversational discourse.
AUTOTUTOR Interface

• AutoTutor is an animated pedagogical agent and it’s interface is comprised of four features:
  – A two-dimensional talking head,
  – A text box for typed student input
  – A text box that displays the problem/question being discussed
  – A graphics box that displays pictures and animations that are related to the topic at hand.
AUTOTUTOR
Computer Literacy Example
AUTOTUTOR
Architecture

• AutoTutor is a combination of classical symbolic architectures (e.g., those with propositional representations, conceptual structures, and production rules) and architectures that have multiple software constraints (e.g., neural networks, fuzzy production systems).

• AutoTutor’s major modules include an animated agent, a curriculum script, language analyzers, latent semantic analysis (LSA), and a dialog move generator.
AUTOTUTOR
Instructional Environment

• Instructional Environment in AutoTutor is represented by the Animated Agent, and the Language Analyzers
• It interacts with the Dialog Move Generator and it modifies the expression according to the dialog AutoTutor is designed to simulate the dialog moves of effective, normal human tutors
• AutoTutor produces dialog moves with pedagogical value, and sensitive to learner’s abilities, within a coherent conversational environment
AUTOTUTOR
Tutoring Dialog

• Five-step dialogue frame specific to human tutoring:
  – Step 1: Tutor asks question (or presents problem).
  – Step 2: Learner answers question (or begins to solve problem).
  – Step 3: Tutor gives short immediate feedback on the quality of the answer (or solution).
  – Step 4: Tutor and learner collaboratively improve the quality of the answer.
  – Step 5: Tutor assesses learner’s understanding of the answer.

• AutoTutor does not contain step 5.
AUTOTUTOR
Animated Agent

• The agents for the AutoTutor programs were created in Curious Labs Poser 4 and are controlled by Microsoft Agent.
• Each agent is a three-dimensional embodied character that remains on the screen throughout the entire tutoring session.

- The agent communicates with the learner via synthesized speech, facial expressions, and simple hand gestures.
AUTOTUTOR
Authoring Tools

• The authoring tools enable experts from various disciplines to easily create content that can be used in AutoTutor tutoring sessions.

• Typically, experts have deep knowledge of subject domains but limited technical and programming skills, whereas the designers of learning technologies have advanced technological knowledge but limited domain expertise.

• User-friendly authoring tools ensure high quality tutoring content for students.
AUTOTUTOR
Authoring Tools

• **Case-based help** - a case study replicating the process that teacher would go through to create a curriculum script using the tool. The scenario was created through an analysis of think aloud protocols with actual teachers during the evaluation process.

• Problems and solutions with the terminology, interface, or concepts were used to generate the case study components, which were then incorporated into an overall composite scenario accessible at any time during the authoring process.
AUTOTUTOR
Authoring Tools

• **Point and Query** - a list of questions-answers units accessible from any part of the tool.

• They are context sensitive Frequently Asked Questions, available through a help button.

• **Glossary** – provides precise definitions for terminology in the script authoring process.

• In the authoring tool, certain terms are hyperlinked to a window that gives the definition of the term.
Q1: What are the important words?
A1: The important words are the words included in the ideal answer points that are directly related to the content. They are very specific in terms of the subject area.

Q2: How do I identify the important words?
A2: Important words are equivalent to a vocabulary list given at the beginning of a lesson.

Q3: How many important words are there for each main point?
A3: That is dependent on how specific the main points are. There could be one important word or all the words in the main point could be important. Don’t include conjunctions, articles, etc.
AUTOTUTOR
Expert Model

- Curriculum Script contains all problems and answers for a particular domain, representing the Expert model. For each problem, it has:
  - an ideal answer,
  - expected good answers,
  - misconceptions,
  - anticipated question-answer pairs,
  - a list of important concepts,
  - problem-related dialog moves.
The problem-related dialog moves currently being used by AutoTutor are:

- Hint
- Prompt
- Prompt Completion,
- Pump,
- Assertion

- Summary,
- Misconception
- Verification
- Correction.
AUTOTUTOR
Curriculum Script Sample

- The curriculum script in AutoTutor organizes the topics and content of the tutorial dialog.
- The general structure of the curriculum script:
  - Macrotopics
  - Topics
  - Dialog moves
AUTOTUTOR
Computer Literacy Curriculum Script

3 Macrotopics
- hardware
- operating systems
- internet

12 Topics each

Topic:
- basic concepts
- focal question
- ideal answers, answer aspects
- hints, prompts
- anticipated bad answers
- corrections for bad answers
- a summary
Large, multi-user computers often work on several jobs simultaneously. This is known as concurrent processing. (...) So here's your question.

How does the operating system of a typical computer process several jobs with one CPU?
The OS helps the computer to work on several jobs simultaneously by rapidly switching back and forth between jobs.

How can the OS take advantage of idle time on the job?

The operating system switches between jobs.
The operating system switches rapidly between jobs.

The operating system completes one job at a time and then works on another.

The operating system can work on several jobs at once.
AUTOTUTOR
The Dialog Advancer Network

• The Dialog Advancer Network (DAN) manages the conversation that occurs between a student and AutoTutor.

• The DAN is comprised of a set of customized dialog pathways that are tailored to particular student speech act categories.

• The DAN enables AutoTutor to adapt each dialog move to the preceding student turn and respond appropriately.
Okay, let's begin. Hello, I'm Marco. I'm one of the computer tutors developed by the Tutoring Research Group, at the University of Memphis. I try to help students learn about Physics, by having a conversation with them. Let's get straight to the demonstration. Suppose a runner is running in a straight line at constant speed, and the runner throws a pumpkin straight up. Where will the pumpkin land? Explain.

The pumpkin lands behind the runner.

Anything else?

I don't know.

Okay. What is the horizontal velocity of the pumpkin?

The pumpkin has a constant horizontal velocity.

So? The horizontal velocity of the pumpkin is the same as the?

Could you repeat that?

Once again, the horizontal velocity of the pumpkin is the same as the?

runner.

The same as the runner. Okay. Also, the man and the pumpkin have the same horizontal velocity. What can we say about the horizontal velocity of the pumpkin if there is no force acting on the pumpkin in the horizontal direction?

its constant.

Note that this transcript only presents a small part of the rich information (pauses, intonation, gestures) that AutoTutor provides in the interaction with a student.
AUTOTUTOR
Categories of Dialog Moves

- (1) Positive immediate feedback. "That's right" "Yeah"
- (2) Neutral immediate feedback. "Okay" "Uh-huh"
- (3) Negative immediate feedback. "Not quite" "No"
- (4) Pumping for more information. "Uh-huh" "What else"
- (5) Prompting for specific information. "The primary memories of the CPU are ROM and _____"
- (6) Hinting. "The hard disk can be used for storage" or “What about the hard disk?”
- (7) Elaborating. “CD ROM is another storage medium.”
- (8) Splicing in correct content after a student error. This is a correction.
- (9) Requestioning. "So once again, what is the function of a CPU?"
- (10) Summarizing. "So to recap," <succinct recap of answer to question>
AUTOTUTOR
DAN – General
TUTOR-1: Alright, let's go on. You want to take photos and send them over the Internet. Consider this problem. What hardware do you need to take photos and send them over the Internet?

STUDENT-1: First of all you need to have access to the internet.

TUTOR-2: What else?

STUDENT-2: a camera

TUTOR-3: Yeah. For what type of connection do you need a network card?

STUDENT-3: network card

TUTOR-4: Okay. Now, if you have access to the Internet through a network card, then your connection is ...

STUDENT-4: I don't know.
• TUTOR-5: Is direct. Okay. A network card is needed if you have a direct connection to the Internet. Okay. How does the user get hooked up to the internet?
• STUDENT-5: by a plug.
• TUTOR-6: Once again, a network card is needed if you have a direct connection to the Internet. Okay, what does a modem do?
• STUDENT-6: connects network card and internet
• TUTOR-7: Okay. Try this. A modem is required when there is ...
• STUDENT-7: a connection TUTOR-8: a dial up connection. . . . 
• TUTOR-30: Let's review. To send your photos on the Internet, you need either a digital camera or a regular camera to take the photos. If you use a regular camera, you need a scanner to scan them onto a computer disk. If you have a direct connection to the Internet, then you need a network card. A modem is needed if you have a dial up connection.
AUTOTUTOR
DAN - Pathway

AutoTutor: Well, where is most of the information you type in temporarily stored?

Student Turn N

Adaption

Select Short Feedback

Tutor Selects Dialog Move

Advancer State

Select Summary

Asks next tutor topic question

Student Turn N + 1

AutoTutor: Right! In RAM.

AutoTutor: Let's review, after you enter information, it is sent to the CPU. The CPU carries out the instructions on the data.

AutoTutor: Okay.

AutoTutor: How does the OS of a typical computer process several jobs simultaneously with only one CPU?
AUTOTUTOR
Language Analyzers

• Language analyzers that are based on recent advances in computational linguistics.

• The purpose of the language analyzers is to improve the conversational smoothness of the system as well as to enhance mixed-initiative dialog.

• The language analyzers include:
  – A word and punctuation segmenter,
  – A syntactic class identifier
  – A speech act classification
AUTOTUTOR
Language Analyzers - Structure

Student's contribution

Word Segmenter

Syntactic Class Identifier

Speech Act Classification
- Assertion
- WH-question
- Yes-/No- question
- Directive
- Short Response

Latent Semantic Analysis
AUTOTUTOR

Dialog Moves Production

• PUMP
  (1) IF [topic coverage = LOW or MEDIUM after learner’s first Assertion] THEN [select PUMP]
  (2) IF [match with good answer bag = MEDIUM or HIGH & topic coverage = LOW or MEDIUM] THEN [select PUMP]
• POSITIVE PUMP
  (3) IF [topic coverage = HIGH after learner’s first Assertion] THEN [select POSITIVE PUMP]
• SPLICE
  (4) IF [student ability = LOW or MEDIUM & student verbosity = LOW or MEDIUM & topic coverage = LOW or MEDIUM & match with bad answer bag = HIGH] THEN [select SPLICE]
• PROMPT
  (5) IF [student verbosity = LOW & topic coverage = LOW or MEDIUM] THEN [select PROMPT]
• HINT
  (6) IF [student ability = MEDIUM or HIGH & match with good answer bag = LOW] THEN [select HINT]
  (7) IF [student ability = LOW & student verbosity = HIGH & match with good answer bag = LOW] THEN [select HINT]
• SUMMARY
  (8) IF [topic coverage = HIGH or number of turns = HIGH] THEN [select SUMMARY]
• ELABORATIONS
  (9) IF [topic coverage = MEDIUM or SOMEWHAT HIGH] THEN [select ELABORATE]
• POSITIVE FEEDBACK
  (10) IF [match with good answer bag = HIGH or VERY HIGH] THEN [select POSITIVE FEEDBACK]
• NEGATIVE FEEDBACK
  (11) IF [match with bad answer bag = HIGH or VERY HIGH & topic coverage = MEDIUM or HIGH] THEN [select NEGATIVE FEEDBACK]
• NEUTRAL FEEDBACK
  (12) IF [match with good answer bag = MEDIUM or SOMEWHAT HIGH] THEN [select POSITIVE NEUTRAL FEEDBACK]
  (13) IF [match with bad answer bag = SOMEWHAT HIGH] THEN [select NEGATIVE NEUTRAL FEEDBACK]
• NEUTRAL FEEDBACK
  (14) IF [match with bad answer bag = HIGH or VERY HIGH & topic coverage = LOW] THEN [select NEGATIVE NEUTRAL FEEDBACK]
  (15) IF [match with good answer bag = LOW or MEDIUM] THEN [select NEUTRAL FEEDBACK]
AUTOTUTOR Interactions

• How Does AUTOTUTOR Handle the Student’s Initial Answer to the Question?
• How Does AUTOTUTOR Get the Learner to Do the Talking?
• How Does AUTOTUTOR Know Whether a Student Has Covered an Expectation?
• How Does AUTOTUTOR Select the Next Expectation to Cover?
• How Does AUTOTUTOR Give Feedback to the Student?
• How Does AUTOTUTOR Handle Mixed-Initiative Dialogue?
AUTOTUTOR
Computer Literacy Example

- TUTOR-1: Alright, let's go on. You want to take photos and send them over the Internet. Consider this problem. What hardware do you need to take photos and send them over the Internet?
- STUDENT-1: First of all you need to have access to the internet.
- TUTOR-2: What else?
- STUDENT-2: a camera
- TUTOR-3: Yeah. For what type of connection do you need a network card?
- STUDENT-3: network card
- TUTOR-4: Okay. Now, if you have access to the Internet through a network card, then your connection is ...
- STUDENT-4: I don't know.
AUTOTUTOR
Computer Literacy Example

- TUTOR-5: Is direct. Okay. A network card is needed if you have a direct connection to the Internet. Okay. How does the user get hooked up to the internet?
- STUDENT-5: by a plug.
- TUTOR-6: Once again, a network card is needed if you have a direct connection to the Internet. Okay, what does a modem do?
- STUDENT-6: connects network card and internet
- TUTOR-7: Okay. Try this. A modem is required when there is ...
- STUDENT-7: a connection
- TUTOR-8: a dial up connection
- ....
- TUTOR-30: Let's review. To send your photos on the Internet, you need either a digital camera or a regular camera to take the photos. If you use a regular camera, you need a scanner to scan them onto a computer disk. If you have a direct connection to the Internet, then you need a network card. A modem is needed if you have a dial up connection.
AUTOTUTOR

Handle the Student’s Initial Answer

• After AUTOTUTOR asks the question in the tutor 1 turn, the student gives an initial answer in the student-1 turn. The answer is very incomplete. A complete answer would include all the points in the summary at the final turn (tutor-30).

• What does AUTOTUTOR do with this incomplete student contribution? AUTOTUTOR doesn’t simply grade the answer (for example, good, bad, incomplete, a quantitative score) as many conventional tutoring systems do.

• AUTOTUTOR also stimulates a multturn conversation that is designed to extract more information from the student and get the student to articulate pieces of the answer.

• AUTOTUTOR attempts to get the student to do the talking and explores what the student knows.
Get the Learner to Do the Talking?

• AUTOTUTOR has a number of dialogue moves to get the learner to do the talking. For starters, there are open-ended pumps that encourage the student to say more, such as What else? In the tutor-2 turn.

• *Pumps* are frequent dialogue moves after the student gives an initial answer, just as is the case with human tutors. The tutor pumps the learner for what the learner knows before drilling down to specific pieces of an answer.

• After the student is pumped for information, AUTOTUTOR selects a piece of information to focus on. Both human tutors and AUTOTUTOR have a set of expectations about what should be included in the answer. What they do is manage the multiturn dialogue to cover these expected answers.
AUTOTUTOR
Control Issues

• Expectation 1: You need a digital camera or regular camera to take the photos.
• Expectation 2: If you use a regular camera, you need to scan the pictures onto the computer disk with a scanner.
• Expectation 3: A network card is needed if you have a direct connection to the internet.
• Expectation 4: A modem is needed if you have a dial-up connection.
AUTOTUTOR
Control Issues

• The dialogue moves vary in directness and information content.

• The most indirect dialogue moves are hints - articulated in the form of questions, designed to lead the learner to construct the expected information

• The most direct dialogue moves are:
  – Assertions - articulate the expected information
  – Prompts - try to get the learner to produce a single word in the expectation
1.1 **AutoTutor** (Prompt): See if you can get this. Computers can be used to analyze data by finding averages, standard deviations, or performing many other kinds of . . .?

1.2 **Student to lab assistant**: [Oh, just answer the question? Um. So, am I still supposed to answer that question (pointing to the main topic question that remains on the monitor)?]

1.3 **Student types**: i don't know

1.4 **AutoTutor** (Prompt Response): kinds of statistics.

1.5 **Student to lab assistant**: [Okay, this is really confusing. I don't understand what the question is supposed to be.]

1.6 **Student types**: i don't know
Excerpt 2
• 2.1 *AutoTutor* (Assertion): Also, a computer has a monitor.
• 2.2 *Student to lab assistant*: [I just answer yes or no?]
• 2.3 *Student types*: yes

Excerpt 3
• 3.1 *AutoTutor* (Assertion): And, the CPU uses RAM as a short-term memory.
• 3.2 *Student to lab assistant*: [Should I, like, I can just say how I think he's unintelligible?]
• 3.3 *Student types*: i am not sure
• 3.4 *AutoTutor* (Neutral Feedback): Uh-huh
• 3.5 *AutoTutor* (Assertion): Also, RAM's contents disappear when the computer is turned off.

• 3.6 *Student types*: what?

• 3.7 *AutoTutor* (Neutral Feedback): Uh-huh

• 3.8 *AutoTutor* (Hint): So, how does the CPU get information from RAM?

• 3.9 *Student types*: I am not sure

• 3.10 *AutoTutor* (Neutral Feedback): Uh-huh

• 3.11 *AutoTutor* (Summary): So, RAM is primary memory storage and is changeable. It is used by the CPU as short-term memory.
4.1 *AutoTutor* (Prompt): Now, what kind of architecture allows a RISC processor to execute multiple instructions simultaneously?

4.2 *Student types*: simple

4.3 *AutoTutor* (Prompt Response): A superscalar architecture.

4.4 *AutoTutor* (Dialog Advancer): Okay,

4.5 *AutoTutor* (Elaborate): And, RISC processors have fewer bottlenecks that complicate processing.

4.6 *AutoTutor* (Dialog Advancer) Okay,

4.7 *AutoTutor* (Summary): Now to wrap up, RISC processors are faster and more efficient than CISC processors (*AutoTutor* continues to summarize this topic).
• 5.1 *AutoTutor* (Prompt): Well, where is most of the information you type in temporarily stored?
• 5.2 *Student types*: RAM
• 5.3 *AutoTutor* (Positive Feedback): Right!
• 5.4 *AutoTutor* (Prompt Response): In RAM,
• 5.5 *AutoTutor* (Dialog Advancer): Moving on,
• 5.6 *AutoTutor* (Elaborate): The CPU handles one instruction at a time.
• 5.7 *AutoTutor* (Dialog Advancer): Okay,
• 5.8 *AutoTutor* (Summary): Let’s review, after you enter information, it is sent to the CPU. The CPU carries out the instructions on the data. (*AutoTutor* continues to summarize this topic).
AUTOTUTOR Tasks

Student Has Covered an Expectation?

• AUTOTUTOR attempts to "comprehend" the student input by segmenting the contributions into speech acts and matching the student's speech acts to the expectations.

• Latent semantic analysis (LSA) is used to compute these matches.

• LSA provides the foundation for grading essays, even essays that are not well formed grammatically, semantically, and rhetorically.
Latent Semantic Analysis (LSA)

- Latent Semantic Analysis (LSA) is a statistical technique that compresses a large corpus texts into a space of 100 to 500 dimensions.

- AutoTutor uses LSA to compare student contributions to expected answer units in the curriculum script.

- AutoTutor has successfully used LSA as the backbone for assessing the quality of student assertions, based on matches to good answers and anticipated bad answers in the curriculum script.
AUTOTUTOR
Language Analyzers - Functions

• Language modules analyze the words in the messages that the learner types into the keyboard during a particular conversational turn – using a 10,000 words lexicon
• Each lexical entry specifies its alternative syntactic classes and frequency of usage in the English language. For example, “program” is either a noun, verb, or adjective.
• Each word that the learner enters is matched to the appropriate entry in the lexicon in order to fetch the alternative syntactic classes and word frequency values.
There is also an LSA vector for each word.

A neural network is used to segment and classify the learner’s content within a turn into speech acts.

The neural network assigns the correct syntactic class to word W, taking into consideration the syntactic classes of the preceding word (W-1) and subsequent word (W+1).
• AutoTutor is capable of:
  – segmenting the input into a sequence of words and punctuation marks with 99%+ accuracy
  – assigning alternative syntactic classes to words with 97% accuracy
  – assigning the correct syntactic class to a word (based on context) with 93% accuracy

• Autotutor uses the learner’s Assertions to assess the quality of learner contributions.
AUTOTUTOR

Student Has Covered an Expectation?

• When the tutor’s expectation $E$ is compared with the learner’s speech act $A$, a cosine match score is computed that varies from 0 (no match) to 1.0 (perfect match).

• AUTOTUTOR considers each combination of speech acts that the learner makes during the evolution of an answer to a major question; the value of the highest cosine match is used when computing whether the student covers expectation $E$. 
AUTOTUTOR

Select the Next Expectation to Cover?

• AUTOTUTOR uses LSA in conjunction with various criteria when deciding which expectation to cover next.

• After each student turn, AUTOTUTOR updates the LSA score for each of the four expectations listed earlier. An expectation is considered covered if it meets or exceeds some threshold value.

• One selection criterion uses the zone of proximal development to select the next expectation, which is the highest LSA score that is below threshold.
AUTOTUTOR
Select the Next Expectation to Cover?

• A second criterion uses coherence, the expectation that has the highest LSA overlap with the previous expectation that was covered.
• Other criteria that are currently being implemented are preconditions and pivotal expectations.
• AUTOTUTOR will decide to cover a new expectation in a fashion that both blends into the conversation and that advances the agenda in an optimal way.
• AUTOTUTOR generates a summary after all the expectations are covered (for example, the tutor-30 turn).
AUTOTUTOR
LSA - Functionality

• The K-dimensional space is used when evaluating the relevance or similarity between any two bags of words, X and Y.
• The relevance or similarity value varies from 0 to 1
• In most applications of LSA, a geometric cosine is used to evaluate the match between the K-dimensional vector for one bag of words and the vector for the other bag of words.
• From the present standpoint, one bag of words is the set of Assertions within turn T.
• The other bag of words is the content of the curriculum script associated with a particular topic, i.e., good answer aspects and the bad answers.
each $A_i$ has coverage metric between 0 and 1 (computed by LSA, updated with each assertion)

• each $A_i$ covered if coverage metric above a threshold
AUTOTUTOR

LSA - Example

Coverage values

\( A_2 \) is covered (above threshold)

Threshold

AutoTutor-1: all contributions count
AutoTutor-2: only student contributions are considered

\( A_5 \) has highest subthreshold value - selected as next GAA to be covered
LSA Values

• LSA values are geometric cosines that range between 0 and 1, where higher values indicate a greater conceptual match.

• The generic values specified in the production rules (e.g., LOW, MEDIUM, HIGH) correspond to a range of arbitrarily determined LSA values.

• In AutoTutor, a HIGH value typically corresponds to LSA values that range between .5 and 1.0, whereas the MEDIUM value typically corresponds to values between .25 and .75.
AUTOTUTOR
Give Feedback to the Student?

• There are three levels of feedback:
  – backchannel feedback that acknowledges the learner's input.
  – evaluative pedagogical feedback on the learner's previous turn based on the LSA values of the learner's speech acts.
  – corrective feedback that repairs bugs and misconceptions that learners articulate.
AUTOTUTOR

Give Feedback to the Student?

• Backchannel feedback: AUTOTUTOR periodically nods and says uh-huh after learners type in important nouns but is not differentially sensitive to the correctness of the student's nouns.

• The backchannel feedback occurs online as the learner types in the words of the turn. Learners feel that they have an impact on AUTOTUTOR when they get feedback at this fine-grain level
AUTOTUTOR
Give Feedback to the Student?

• Evaluative pedagogical feedback - The facial expressions and intonation convey different levels of feedback, such as:
  – negative (for example, not really while head shakes)
  – neutral negative (okay with a skeptical look)
  – neutral positive (okay at a moderate nod rate)
  – positive (right with a fast head nod).
AUTOTUTOR
Give Feedback to the Student?

• Corrective feedback - The bugs and their corrections need to be anticipated ahead of time in AUTOTUTOR'S curriculum script.

• An expert tutor often has canned routines for handling the particular errors that students make. AUTOTUTOR currently splices in correct information after these errors occur.

• Sometimes student errors are ignored because it evaluates student input by matching it to what it knows in the curriculum script, not interpreting
AUTOTUTOR
Handle Mixed-Initiative Dialogue?

01 AUTOTUTOR: When you turn on the computer, how is the operating system first activated and loaded into RAM?
02 STUDENT: When the computer is turned on a file is automatically booted up.
03 AUTOTUTOR: Anything else?
04 STUDENT: This file contains the operating system.
05 AUTOTUTOR: What else?
06 STUDENT: The operating system is then read into RAM and activated.
07 AUTOTUTOR: Right. The CPU executes instructions in ROM, which helps the operating system to start.
08 STUDENT: After being read from ROM the OS is activated.
09 AUTOTUTOR: Right. Let’s review what we’ve just gone over. The operating system must first be stored on the hard disk. When you turn on the computer, the CPU executes instructions stored in ROM. These instructions help boot the operating system. The operating system is then loaded from the hard disk into RAM.
Handle Mixed-Initiative Dialogue?

- AutoTutor first asks a *how*-question in turn 1.
- Then the student provides answer information in turns 2, 4, 6, and 8.
- These answers are encouraged by AutoTutor’s *pumps* in turns 3 and 5 (i.e., “Anything else?”, “What else?”). AutoTutor delivers these pumps in an effort to encourage the student to actively construct knowledge.
- Thus, instead of being an *information delivery system* Teaching Tactics and Dialog in AutoTutor that bombards the student with a large volume of information, AutoTutor is a *discourse prosthesis* that attempts to get the student to do the talking and that explores what the student knows.
AUTOTUTOR
Handle Mixed-Initiative Dialogue?

• AutoTutor adopts a constructivist theoretical view that a key feature of effective learning lies in assisting students in actively constructing subjective explanations and elaborations of the material, as students answer questions and solve problems that require deep reasoning.

• At the same time, however, the students need to answer enough questions and to solve enough problems for them to understand the constraints of the domain knowledge.

• AutoTutor brings the student back on track by supplying cues and clues that lead to the evolution of a complete answer to the question. These clues include hints, prompts for the student to fill in a word or phrase, and assertions that fill in missing ideas.
The student had forgotten about the role of ROM in launching the operating system, so AutoTutor brings up ROM in turn 7. The student builds on this suggestion in turn 8. At that point, the important pieces of a good complete answer have been covered, so AutoTutor summarizes the answer in turn 9. AutoTutor periodically gives positive immediate feedback after the student contributions (i.e., “right.”). This feedback is not only motivating, but creates the impression that AutoTutor is listening to what the student is communicating. These characteristics of a tutorial exchange are quite similar to discourse patterns in normal tutoring between humans.
AUTOTUTOR
Evaluation

• Tests on AutoTutor as effective tutor and conversational partner in three evaluation cycles
• The purpose of the evaluation cycles was to identify and correct particular dialog move problems before AutoTutor’s debut with human learners.
• After each evaluation cycle, the curriculum script, the fuzzy production rules, and the LSA parameter thresholds were revised to enhance AutoTutor’s overall performance.
• Several virtual students were created to emulate human students of varying ability and verbosity levels.
AUTOTUTOR
Evaluation

• 1. Good verbose student. The first 5 turns of this virtual student had 2 or 3 Assertions that human experts had rated as good Assertions from the human sample. The student is regarded as verbose because the student has 2 or 3 Assertions within one turn, which is more than the average number of Assertions per turn in human tutoring.

• 2. Good succinct student. The first 5 turns of this virtual student had 1 Assertion that human experts had rated as a good Assertion.

• 3. Vague student. The first 5 turns of this virtual student had an Assertion that had been rated as vague (neither good nor bad) by the human experts.
• 4. Erroneous student. The first 5 turns of this virtual student had an Assertion that contained a misconception or bug according to human experts.

• 5. Mute student. The first 5 turns of this virtual student had semantically depleted content, such as “Well”, “Okay”, “I see”, and “Uh”. Person, Graesser, Kreuz, Pomeroy, and the Tutoring Research Group

• 6. Good coherent student. The first 5 turns of this virtual student had 1 Assertion that had been rated as good. All of the Assertions in the first 5 turns for a particular topic were provided by one human student.

• 7. Monte Carlo student. The first 5 turns of this virtual student were generated in a Monte Carlo fashion to simulate the variability of student Assertion quality that typically occurs human tutoring sessions.
AUTOTUTOR Evaluation

• Four judges rated the quality of AutoTutor’s dialog moves on two holistic dimensions:
  – pedagogical effectiveness (PE)
  – conversational appropriateness (CA).

• Two judges were assigned to each dimension.

• For each AutoTutor dialog move, the PE judges considered:
  – (1) whether the dialog was pedagogically effective,
  – (2) whether the dialog move was reasonable for a normal human tutor.
AUTOTUTOR

Evaluation

• The CA judges considered several factors relevant to conversation in their holistic rating of each AutoTutor dialog move: politeness norms along with the Gricean maxims of quality, quantity, relevance, and manner.

• Both PE and CA were rated on a six-point scale, where 1 reflected a very low quality rating and 6 reflected a very high quality rating. Inter-judge reliability measures were computed for both pairs of judges.

• Results indicated significant reliability between judges for both dimensions (Cronbach’s alpha = .94 for PE and .89 for CA).
6 human tutors were asked what they would say at these 144 points.

Transcripts of AutoTutor-1's dialog moves.

144 Tutor Moves from Dialogs between Students and AutoTutor-1.

36 computer literacy students discriminated: AutoTutor or Human Tutor?
Outcome: discrimination score of -.08.
AUTOTUTOR
TRG conclusions

• “Impressive” outcome supported claim that AutoTutor is a good simulation of human tutors.
• Attempts to comprehend the student input.
• “Almost as good as an expert in computer literacy.”
AUTOTUTOR
Emotional Responses

• Students initially amused by the talking head – but amusement wears off in a few minutes.
• Trouble in understanding the synthesized speech (some students).
• Inappropriate speech acts irritate students (only minority).
• Sufficiently engaging to complete the tutorial sessions.
AUTOTUTOR
Conclusions

• Strengths
  – not purely domain-specific
  – easy creation of curriculum script (no programming skills needed)
  – robust behaviour

• Weaknesses
  – shallow understanding only
  – performance largely depends on Curriculum Script
References

- Eric Horvitz, “Principles of Mixed-Initiative User Interfaces”, Microsoft Research, Redmond, WA.
- AUTOTUTOR website: www.autotutor.org
- Arthur C. Graesser1, Xiangen Hu1, Suresh Susarla1, Derek Harter1, Natalie Person2, Max Louwerse1, Brent Olde1, and the Tutoring Research Group1 AutoTutor: “An Intelligent Tutor and Conversational Tutoring Scaffold”, http://www-2.cs.cmu.edu/~aleven/AIED2001WS/Graesser.pdf
References

