Intelligent Tutoring Systems: New Challenges and Directions

Cristina Conati
Department of Computer Science
University of British Columbia
2 sigma effect (Bloom, 1984):

- Average student’s achievement with a personal human tutor better than 98% of classroom students

Intelligent Tutoring Systems (ITS)

- Interdisciplinary field

- Aiming to
  - Create computer-based tools that support individual learners
  - By autonomously and intelligently adapting to their specific needs
Outline

- Background and definitions
- Achievements and new research directions
- Sample Projects
Precursors of ITS

- Computer-Assisted Instruction (CAI) systems

Artificial Intelligence

Cognitive Science

Human-Computer Interaction

Education
Precursors of ITS

- Computer-Assisted Instruction (CAI) systems

Shute and Potska 1996, Handbook of Research on Educational Communications and Technology
CAI systems (cont.)

- All branching in the program to be pre-defined
  - Sequencing of topics and exercises
  - All relevant student answers
  - All feedback actions

- Student’s solution process is not taken into account, only final answers
  - No information on the reasons for the student behavior

- Fine for drill-and-practice in simple domains such as basic math operations
  - Unmanageable for more complex domains and pedagogy (e.g., support problem solving in physics)
A 2000-kg car in neutral at the top of a 20-degree inclined driveway 20 m long slips its parking brake and rolls down. Assume that the driveway is frictionless.

At what speed will it hit the garage door?

Answer: 

\[ v_f \]
Good Human Tutors..

- Can provide more flexible and comprehensive support to learning
  - Recognize a large variety of student’s behaviors
  - Diagnose student’s understanding (and other relevant states)
  - Provide adequate tailored interventions at different stages of the interaction
Intelligent Tutoring Systems (ITS)

Artificial Intelligence
- **Represent** knowledge and processes relevant for effective tutoring
- **Reason** to select effective tutorial actions
- **Learn** from experience

Cognitive Science

Education

Human-Computer Interaction
Ideal ITS

- Pedagogical model
  - Teaching strategies
  - Remediation
  - Curriculum

- Communication model

- Domain Model
  - Concepts
  - Principles...

- Solution Generator
  - Computer solution (solution step)

- Tutor
  - Tutorial Action
    - Select activity
    - Hints
    - Feedback
    - Corrections
    - Etc.

- Interface

- Student Modeler
  - Student model
    - Knowledge, Goals, Beliefs...

- Student solution (solution step)
Achievements

- In the last 20 years, there have been many successful initiatives in devising Intelligent Tutoring Systems (Woolf 2009, Building Intelligent Interactive Tutors, Morgan Kaufman)
  - Including CanergieLearning, a company that commercializes ITS for Math in hundreds of high schools in the USA
How much learning improvement?

Most sophisticated CAI: 0.5 $\sigma$ (Dodds & Fletcher, 2004)
However…

- Mainly ITS that provide individualized support to problem solving through tutor-lead interaction (*coached problem solving*)
  - Well defined problem solutions => guidance on problem solving steps
  - Clear definition of correctness => basis for feedback
Coached problem solving

- Example: Andes, tutoring system for Newtonian physics

Beyond Coached Problem Solving

- Coached problem solving is a very important component of learning.

- Other forms of instruction, however, can help learners acquire the target skills and abilities:
  - At different stages of the learning process
  - For learners with specific needs and preferences
Key Trends in ITS

Adaptive Open Learning Environments
- Support learning via free exploration of *virtual worlds*, *interactive simulations* and *educational games*.

Collaborative Learning Environments
- Adaptive scaffolding of effective group-based learning.

Affective Tutors
- Understand and react to learners' emotions.

Meta-Cognitive Tutors
- Scaffold acquisition of learning and reasoning (*meta-cognitive*) skills.
Challenges

- Activities more open-ended and less well-defined than pure problem solving
  - No clear definition of correct/successful behavior
- Higher-level user states (meta-cognitive, affective)
  - difficult to assess unobtrusively from interaction events
- High level of uncertainty
Key Trends in ITS

Adaptive Open Learning Environments
- Support learning via free interaction with virtual worlds, interactive simulations and educational games

Collaborative Learning Environments
- Adaptive scaffolding of effective group-based learning

Affective Tutors
- Understand and react to learners emotions

Meta-Cognitive Tutors
- Scaffold acquisition of learning and reasoning (meta-cognitive) skills
Our Approach

- Student models based on formal methods for probabilistic reasoning
  - Bayesian networks and extensions

- Decision theoretic approach to tutorial action selection

- Increase the bandwidth through innovative input devices:
  - e.g. eye-tracking and physiological sensors
Outline

- Background and definitions
- Achievements and Trends
- Sample Projects
Adaptive Open Learning Environments
- Support learning via free exploration of virtual worlds, interactive simulations and educational games

Affective Tutors
- Understand and react to learners' emotions

Decision theoretic support for Analogical Problem Solving

Meta-Cognitive Tutors
- Scaffold acquisition of learning and reasoning (meta-cognitive) skills
Students find it helpful to refer to examples in the early stages of problem solving (e.g. Reed and Bolstad, 1991)

**Analogical Problem Solving (APS)**

But the effectiveness of APS is mediated by two meta-cognitive skills (Vanlehn 1998)
A person pulls a 9 kg crate up a ramp inclined 30 degrees to the horizontal. The pulling force is applied at an angle of 30 degrees CCW from the horizontal, with a magnitude of 100N. The crate's acceleration is not known. Find the normal force on the crate.

We answer this question by using Newton's Second Law.

First, we choose the crate as the body, using the body tool.

Next, we define the body's properties:
- The mass of crate is 9 kg
- The crate's acceleration is unknown

Next, we find all the forces acting on the crate using the force tool.

The weight force \( W \) on the crate is due to the Earth.
- It is oriented at 270 degrees
- It's magnitude is: \( W = \text{gravity} \times \text{mass} \), which is 9.8 m/s²

The normal force \( N \) on the crate is due to the floor.
- It is oriented at 120 degrees
- The pulling force applied to the crate

To find the normal force on the crate, we apply Newton's Second Law along the \( y \)-axis:

\[
\text{net}_y = m \times a_y
\]

\[
N_y + W_y + P_y = m \times a_y
\]

To solve the above equation, we need to decompose all the forces and accelerations into their \( y \) components:

\[
N_y = N
\]

\[
W_y = -W \times \sin(80)\]

\[
P_y = 0
\]

\[
a_y = 0
\]

Substituting these into our Newton's equation:

\[
N_y - W \times \sin(30) = m \times a_y
\]
1. Min-Analogy
   - minimize copying from examples

2. Self-explanation: tendency to elaborate and clarify to oneself given instructional material (Chi et al, ’89)
   - can be used to *learn* new domain principles
Impact of Student Characteristics

- Unfortunately, some students lack these skills (e.g. Vanlehn 1999)
  - Maximize copying
  - Don’t learn new domain principles via SE

- Furthermore, lack of domain expertise leads to selection of inappropriate examples (e.g. Novick 1988)
Impact of Problem/Example Similarity

| Low Problem-Solving Success & Learning | X | X |
| High Problem-Solving Success & Learning | ✓ | X |

- These effects are mediated by student existing knowledge and meta-cognitive skills, e.g.
  - Student with tendency for min-analogy and SE may still benefit from a very similar example
Adaptive support for Example Studying

The EA (Example Analogy) Coach
(Muldner and Conati, User Modeling 2005)

- Suggests examples that aid *both* problem solving success and learning
  - By supporting min-analogy and SE

- Provides interface scaffolding to help learners use them effectively *(demo)*
Dynamic Bayesian network
- evolution of student knowledge and studying behaviour given student interface actions

Example-Selection Mechanism

Student Model

Expected Utility Calculation

Solver

Solution Graph

Knowledge Base

Problem Specification

Coach

Interface

Example Pool

Problem Pool

how intermediate solution steps derive from physics rules and previous steps
Decision Theoretic Approach to Example Selection

Example Solution

Steps Similarity

Problem Solution

Probabilistic Student Model
- Knowledge of physics rules
- Tendency for SE and Min-analogy

Simulation of problem solving

Prediction of learning & problem solving success

Multi-attribute Utility Function gives expected utility of example based on these predictions

- Done for every example known by the EA-Coach
  - Select example with Maximum Expected Utility
Evaluation
(Muldner and Conati, IJCAI 2007)

- 16 participants solved problems with the EA-Coach
- Within subjects design

Conditions

- Adaptive: examples selected via the decision-theoretic mechanism
- Static: example pre-selected for each problem as the most similar in the available pool
  - standard approach currently used in ITS (e.g., Weber 1996)
Results

- **Learning**
  
  - Adaptive condition generated significantly more self-explanations and fewer copy episodes: *better learning*

- At no cost for *problem solving success*
  
  - No significant differences in problem solving success between conditions
  
  - The adaptive condition on average made more errors and took longer to solve problems

Bi-product of learning
Discussion

- Encouraging evidence that our proposed decision-theoretic approach to example selection triggers the desired behaviors

- Future work
  - More proactive adaptive interventions
  - *Eye-tracking* instead of masking interface to capture student reasoning
Eye Tracking and Self-explanation

- We have already investigated the value of eye-tracking information to monitor student self-explanation.
- Domain: *interactive simulations* for understanding mathematical functions.
Sample Activity
User Model

(Bunt, Muldner and Conati, ITS2004; Merten and Conati, Knowledge Based Systems 2007)

Interface actions
Input from eye-tracker

User Model (Dynamic Bayesian Network)

- Number and coverage of student actions
- Self-explanation of action outcomes
  - Time between actions
  - Gaze Shifts in Plot Unit

Learning
\[ f(x) = -3(x+1.7)^2 + 1.9 \]
Results on Accuracy

Accuracy on SE  Accuracy on Learning

- No SE
- SE (Time)
- SE (Time + Gaze)
Discussion

- Evidence that eye-tracking can support real-time modeling of user reasoning processes

- Future work
  - Data mining of eye-tracking + interface actions to uncover other relevant attention patterns (Amershi and Conati, IUI 2007)
  - Devise and test adaptive interventions to trigger self-explanation and exploration
Adaptive Open Learning Environments
-Support learning via free exploration of virtual worlds, interactive simulations and educational games

Affective Tutors
-Understand and react to learners emotions

Modeling Student Affect in Educational Games

Meta-Cognitive Tutors
- Scaffold acquisition of learning and reasoning (meta-cognitive) skills
Educational Games

- Educational systems designed to teach via game-like activities

- Usually generate high level of emotional engagement and motivation.

- Still little evidence on pedagogical effectiveness (e.g. Vogel 2004, Van Eck 2007)
  - Often possible to play well without reasoning about the target instructional domain
Example: The Prime Climb Educational Game

Designed by EGEMS group at UBC to teach number factorization to students in 6th and 7th grade (11 and 12 year old)
Our Solution

Emotionally Intelligent Pedagogical Agents that

- Monitor how students learn from a game
- Generate tailored interventions to help students learn…
- …while maintaining a high level of student emotional engagement

Crucial to model student affect in addition to learning
Initial Prime Climb Pedagogical Agent
(Conati and Zhao, Intelligent User Interfaces 2004)

- Answers to students’ help requests
- Provides unsolicited hints
  - Both after correct and incorrect moves
- Based only on a probabilistic model of student’s knowledge
Hints at Incremental Level of Detail

Think carefully about how to factorize the number you clicked on.

Factors are numbers that multiply to give the number. Look at this example.

The factors of 12 are 2, 3, 4, 6, and 12
because
12 = 1 x 12 (don't include 1)
12 = 2 x 6
12 = 3 x 4

Do you want another hint?
NO YES

OK
What else do we need?

- Prime Climb, with the model of student knowledge and the agent, generated better learning than the basic game (Conati and Zhao IUI 2004).

- But by taking affect into account, the agent could do even better
  - Decide what to do when the student is upset
  - Decide how to use student’s positive states to further improve learning
Long Term Goal

A decision-theoretic *Emotionally Intelligent* Agent for Prime Climb

Model of Student Knowledge

Model of Student Emotional State

Possible actions + predicted effect on learning and affect

Select actions to optimize balance between learning and emotional engagement
How to Assess Emotions?

- Emotions can be assessed by
  - Reasoning about possible causes (i.e. the interface keeps interrupting the user, so she is probably frustrated)
  - Looking at the user’s reactions
But Things are not Always that Easy

- The mapping between emotions, their causes and their effects can be highly ambiguous
  - Different people react differently to similar events
  - How much emotions are shown may depend on culture, personality, current circumstances

- Very hard to build models of user affect
Challenge

- Difficulty of modeling affect in edu-games enhanced by the fact that players often experience
  - Multiple emotions
  - Possibly overlapping
  - Rapidly changing

- For instance:
  - Happy with a successful move but upset with the agent who tells them to reflect about it
  - Ashamed immediately after because of a bad fall
Previous Approaches

- Reduce uncertainty by modeling
  - one relevant emotion in a restricted situation (e.g., Healy and Picard, 2000; Hudlicka and McNeese, 2002)
  - only intensity and valence of emotional arousal (e.g. Prendinger 2005)

- Model longer term, mutually exclusive emotions like boredom, frustration, flow (D’Mello et al 2008, Arroyo et al 2009)

- Not sufficient to react promptly to the more instantaneous, possibly overlapping emotions observed in Prime Climb
Our solution

- Base emotion assessment on information on both causes and effects of emotional reaction

- Probabilistically combined in a Dynamic Bayesian Network
The Prime Climb Affective Model

Game-based Causes

Predictive Assessment

Emotional State

Diagnostic Assessment

Player Reactions

Based on the OCC Theory of Emotions (Ortony Clore and Collins, 1998)
OCC Theory

Defines 22 different emotions are the result of evaluating (appraising) the current circumstances with respect of one’s goals.
The Predictive Part of the Model

Game-based Causes

Predictive Assessment

Emotional State

Infers **player goals** at runtime from **personality traits** and **interaction events**

- Have Fun
- Learn Math
- Succeed by Myself
- Want Help
- Beat Partner

Has information to assess which game states satisfy/dissatisfy the goals

6 of the 22 emotions in the OCC theory

- Joy/Regret toward the game
- Admiration/Reproach toward the agent
- Pride/Shame toward oneself

Conati and MacLaren 2009, J. of User Modeling and User-Adaptive Interaction
DDN for the OCC Theory

All the relations in the model have been defined via empirical studies and observations (Conati and MacLaren 2005)
Evaluation of the Diagnostic Model

- User study with 66 students (10 and 11 year of age) (Conati and Maclaren UMUAI 2009)
  - Pretty good results in capturing emotions towards the game (69.5% accuracy)
  - Not so good for emotions towards the agents, specifically in recognizing regret (47% accuracy)

- Due to the model’s inability to capture goal shifts from Succeed-by-myself to Want-help at difficult times
  - Because goals are modeled as static
Adding Diagnostic Information
(Conati and McLaren, User Modeling 2009)

Game-based Causes

Predictive Assessment

Emotional State

Diagnostic Assessment

Player Reactions
Diagnostic Assessment

- **Long term goal**: integrate multiple detectors: physiological sensors, face and intonation recognition

- **Current focus**: Electromyogram (EMG)
  - Applied on the forehead detects activity in the corrugator muscle
  - greater activity is a reliable indicator of negative affect (e.g., Cacioppo 1993)
    - emotions expressed on demand
EMG signal

Signal peak in correspondence of a frown
The Prime Climb Affective Model

Game-based Causes

Predictive Assessment

• Infers player goals at runtime (e.g., Have Fun, Learn Math, Avoid Falling…)
• Has information to assess which game states satisfy/dissatisfy the goals

Emotional State

Diagnostic Assessment

• Joy/Regret toward the game
• Admiration/Reproach toward the agent
• Pride/Shame toward oneself

Overall Valence

Relevant probabilities learned from data from ad-hoc study

EMG Signal
Evaluation

- Tested with 41 students (from 6th and 7th grade)
  - Periodically reported their emotions during game playing
  - Model predictions compared against these self-reports
Emotional Reports

How do you feel about your game playing?
Very Bad  Bad  Neutral  Good  Very Good

How do you feel about the agent?
Very Bad  Bad  Neutral  Good  Very Good
Evaluation

- Good results in presence of strong, equally-valenced emotions
  - Adding EMG significantly improves accuracy

- Weaker results in presence of subtler or conflicting emotions
Comparison of Predictive and Complete Model

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Accuracy</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Predictive Model</td>
<td>Combined Model</td>
</tr>
<tr>
<td>Joy</td>
<td>74.80</td>
<td>79.10</td>
</tr>
<tr>
<td>Distress</td>
<td>53.48</td>
<td>56.7</td>
</tr>
<tr>
<td>J/D Combined</td>
<td>64.14</td>
<td>67.9</td>
</tr>
<tr>
<td>Admiration</td>
<td>83.49</td>
<td>81.18</td>
</tr>
<tr>
<td>Reproach</td>
<td>39.11</td>
<td>63.02</td>
</tr>
<tr>
<td>A/R Combined</td>
<td>61.3</td>
<td>73.10</td>
</tr>
</tbody>
</table>

- Encouraging evidence that EMG can help model assessment of individual emotions with clear overall valence
Lots of Exciting Future Work

- Add more diagnostic elements to improve model accuracy (e.g., more expression recognition, speech/intonation patterns)
- Integrate model of affect and model of learning
- Create emotionally intelligent agent that takes into account both student affect and learning to decide how to act
- Include longer term emotions (boredom, happiness, frustration)
- Prove that it works better than agent with no affect!
Conclusions

- AI has the potential of having a huge impact on society by affecting how people learn and train
  - Bring benefits of one-to-one tutoring to all
- Successful ITS have already been deployed to support problem solving activities
- The benefits of AI in education and training can go further, e.g.
  - Support for life-long learning via promoting meta-cognition
  - Innovative personalized activities: learning via exploration, learning via game playing
  - Affect-sensitive tutors
- Continuous dialogue with educators and cognitive scientists crucial to do this right
Thanks to…

- Andrea Bunt
- Heather Maclaren
- Cristina Merten
- Kasia Muldner
- David Ternes

And to you all for your kind attention
Evaluation of the Combined Model: Mild/Mixed Valence Data Points

- No improvement over predictive model
- One EMG on forehead not adequate for assessing mild emotions
  - Need to integrate it with other sensors, i.e. heart-rate monitor, EMG on zygomatic muscle
Mixed Valence

Need to improve goal recognition

Unless there is strong evidence coming from the causal component

Positive/Negative Valence forces J/D and A/R to have the same valence.

Evidence from any valence detector forces Valence node to be positive or negative