Toward socially intelligent tutoring systems (ITS)

Kurt VanLehn
Computing, Informatics and Decision Science Engineering
Arizona State University
Outline

• Introduction to ITS
• Why are ITS effective?
  – Feedback and hints on steps (not just answers)
  – Adaptive task selection
• Why are ITS uncommon?
  – Model development is costly
  – ITS disrupt classroom social structure
• Can crowd-sourcing help solve ITS’s problems?
A typical course consists of:

- A sequence of modules, each comprised of a sequence of:
  - Passive media (text, multimedia, lecture)
  - Exercises
  - Discussion
  - Quiz
- Final project
- Final exam
Adding an ITS makes these changes:

- A sequence of modules, each comprised of a sequence of
  - Passive media (text, multimedia, lecture)
  - Exercises
  - Discussion
  - Quiz
- Final project
- Final exam
How does an ITS work?

• It chooses the next activity/task for the student to do based on a model of the student’s current competence, affect and interest.

• It conducts stealth assessment

• If the task is a complex, multi-step activity
  – It understands each step the student makes
  – It can provide hints & feedback on every step
What is the structure of an ITS?

Tool (e.g., editor) emits student steps

Expected steps (correct + misconceived)

Comparison

Hint & feedback generator

Student competences, affect, etc

Next task selector
SLQ-Tutor (Mitrovic & Ohlson, 1999; Addison Wesley)

Problem

List the titles and numbers of all movies that have won at least one Academy Award and have been made in or after 1988.

Step 1

SELECT title, number

Step 2

FROM movies

Step 3

WHERE aawon>1 and year>=1988

Step 4

GROUP BY

Step 5

HAVING

Step 6

ORDER BY

Submit!

Almost there - a few mistakes though. One of them is in the FROM clause. You can correct your query and press 'Submit' again, or try getting some more feedback.

Would you like to have another go?

Feedback

The database that the problem refers to

Schema for the MOVIES Database

The general description of the database is available here. Clicking on the name of a table brings up the table details. Keys in the attribute list are underlined, foreign keys are in italics.

Table Name | Attribute List
---|---
DIRECTOR | number lname fname born died
MOVIE | number title type aonon aawon year critics director
STAR | lname fname number born died city
CUSTOMER | lname fname number address rentals bonus jdate
TAPE | code movie pdate times customer hiredate
STARS_IN | movie star role
Cognitive Algebra I Tutor has multiple tools
www.carnegielearning.com

Problem

Step: Label a column

Step: Fill in a cell

Step: Define an axis

Step: Plot a point

Step: Divide both sides
Andes3 is like power-point, but gives correct (green) vs. incorrect (red) feedback.

ITS can tutor steps in an argument/essay


**Task**

The sun exerts a gravitational force on the earth as the earth moves in its orbit around the sun. Does the earth pull equally on the sun? Explain why.

**Tutor**

No, the sun is much more massive than the earth, so it pulls harder. That is why the earth orbits the sun and not vice versa.

**Tutor-student dialogue**

Student: No, the sun is much more massive than the earth, so it pulls harder. That is why the earth orbits the sun and not vice versa.
A simulation-based ITS where all feedback is delayed until done fighting fire

Step: Isolate a compartment

Debriefing replays steps and discusses

After-action reviews often use a timeline

www.stotler-henke.com

ITS marks learning opportunities with red
ITS can be an non-player character

Learning companion(s) can accompany ITS


---

Concept map editor

ITS

Betty, a teachable agent
What is the structure of an ITS?

- Tool (e.g., editor) emits student steps
- Expected steps (correct + misconceived)
- Comparison
- Hint & feedback generator
- Student competences, affect, etc
- Next task selector
Adding an ITS makes these changes:

• A sequence of modules, each comprised of a sequence of
  – Passive media (text, multimedia, lecture)
  – Exercises
  – Discussion
  – Quiz
• Final project
• Final exam

- dynamically adaptive
- complex, dynamically scaffolded
- unnecessary
Outline

- Introduction to ITS
- Why are ITS effective?
  - Feedback and hints on steps (not just answers)
  - Adaptive task selection
- Why are ITS uncommon?
  - Model development is costly
  - ITS disrupt classroom social structure
- Can crowd-sourcing help solve ITS’s problems?
Micki Chi’s ICAP framework

<table>
<thead>
<tr>
<th>Student engagement activity</th>
<th>e.g., history</th>
<th>e.g., algebra equations</th>
<th>Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive</td>
<td>Listening to a tutor lecture</td>
<td>Watching a tutor’s example</td>
<td>Worst</td>
</tr>
<tr>
<td>Active</td>
<td>Taking notes on tutor’s lecture</td>
<td>Copying an example</td>
<td>OK</td>
</tr>
<tr>
<td>Constructive</td>
<td>Answering questions</td>
<td>Solving a problem</td>
<td>Better</td>
</tr>
<tr>
<td>Interactive</td>
<td>Discussing questions with a peer or tutor</td>
<td>Solving a problem with a peer or tutor</td>
<td>Best</td>
</tr>
</tbody>
</table>

I > C > A > P
Do computer tutors interact frequently enough, compared to humans?

- **No** tutoring (baseline for comparisons)
- **Answer**-based tutoring
  - Hints and feedback on short-answer questions
- **Step**-based tutoring
  - Hints and feedback on steps normally taken when using tool
- **Substep**-based tutoring
  - Tutor can discuss reasoning behind steps
- **Human** tutoring

CAI (computer aided instruction)

ITS
Use Figure 5-20 to determine during what time interval the acceleration is largest and during what time interval the acceleration is smallest.

**Largest acceleration**
- $t = 5.2 \text{ s to } t = 6.0 \text{ s}$
- $t = 10.5 \text{ s to } t = 11.0 \text{ s}$
- $t = 0 \text{ to } t = 0.5 \text{ s}$
- $t = 31 \text{ s to } t = 32 \text{ s}$
Step-based tutoring
Substep-based tutoring


---

Student enters an equation (step)

Tutor asks about reasoning behind the step
A common belief: The more frequent the interaction, the more effective the tutoring.
All possible pairwise comparisons


<table>
<thead>
<tr>
<th>Tutoring type</th>
<th>vs. other tutoring type</th>
<th>Num. of effects</th>
<th>Mean effect</th>
<th>% reliable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer-based</td>
<td>no tutoring</td>
<td>165</td>
<td>0.31</td>
<td>40%</td>
</tr>
<tr>
<td>Step-based</td>
<td>no tutoring</td>
<td>28</td>
<td>0.76</td>
<td>68%</td>
</tr>
<tr>
<td>Substep-based</td>
<td>no tutoring</td>
<td>26</td>
<td>0.40</td>
<td>54%</td>
</tr>
<tr>
<td>Human</td>
<td>no tutoring</td>
<td>10</td>
<td>0.79</td>
<td>80%</td>
</tr>
<tr>
<td>Step-based</td>
<td>answer-based</td>
<td>2</td>
<td>0.40</td>
<td>50%</td>
</tr>
<tr>
<td>Substep-based</td>
<td>step-based</td>
<td>6</td>
<td>0.32</td>
<td>33%</td>
</tr>
<tr>
<td>Human</td>
<td>step-based</td>
<td>1</td>
<td>-0.04</td>
<td>0%</td>
</tr>
<tr>
<td>Substep-based</td>
<td>step-based</td>
<td>11</td>
<td>0.16</td>
<td>0%</td>
</tr>
<tr>
<td>Human</td>
<td>sub-step based</td>
<td>10</td>
<td>0.21</td>
<td>30%</td>
</tr>
<tr>
<td>Human</td>
<td>sub-step based</td>
<td>5</td>
<td>-0.12</td>
<td>0%</td>
</tr>
</tbody>
</table>
Graphing all 10 comparisons: none < answer < step = substep = human

Effect size vs. No tutoring
- vs. Answer-based
- vs. Step-based
- vs. Substep-based

Effectiveness vs. Increasing frequency of interaction
Outline

• Introduction to ITS
• Why are ITS effective?
  – Feedback and hints on steps (not just answers)
  – Adaptive task selection
• Why are ITS uncommon?
  – Model development is costly
  – ITS disrupt classroom social structure
• Can crowd-sourcing help solve ITS’s problems?
The Zone of Proximal Development (ZPD)

(a) Optimal Gameplay Corridor

(b) Zone of Proximal Development
How can we keep students in their ZPDs?

• “Mastery learning” means continuing to work on a module until you have mastered it.

• With an ITS, traditional mastery tests are replaced by the stealth assessment
Mastery learning implemented by the red path below

Tool (e.g., editor) emits student steps

Expected steps (correct + misconceived)

Comparison

Hint & feedback generator

Student competences, affect, etc

Next task selector
Students may be allowed to work at own pace *(Quantum’s accounting tutor)*

<table>
<thead>
<tr>
<th></th>
<th>Jim</th>
<th>Sue</th>
<th>Juan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Borrowing money from bank</strong></td>
<td>In Progress</td>
<td>In Progress</td>
<td>Mastered</td>
</tr>
<tr>
<td><strong>Collecting on receivables</strong></td>
<td>Mastered</td>
<td>Mastered</td>
<td>Mastered</td>
</tr>
<tr>
<td><strong>Events that aren’t transactions</strong></td>
<td>In Progress</td>
<td>In Progress</td>
<td>Mastered</td>
</tr>
<tr>
<td><strong>Investments for stock</strong></td>
<td>In Progress</td>
<td>Not Attempted</td>
<td>In Progress</td>
</tr>
<tr>
<td><strong>Involving more than two accounts</strong></td>
<td>In Progress</td>
<td>In Progress</td>
<td>Not Attempted</td>
</tr>
<tr>
<td><strong>Paying dividends</strong></td>
<td>Mastered</td>
<td>Mastered</td>
<td>Mastered</td>
</tr>
<tr>
<td><strong>Paying expenses</strong></td>
<td>Not Attempted</td>
<td>Not Attempted</td>
<td>Not Attempted</td>
</tr>
<tr>
<td><strong>Payments to creditors</strong></td>
<td>Mastered</td>
<td>Mastered</td>
<td>Mastered</td>
</tr>
<tr>
<td><strong>Prepaying expenses</strong></td>
<td>In Progress</td>
<td>Mastered</td>
<td>In Progress</td>
</tr>
<tr>
<td><strong>Providing services to customers</strong></td>
<td>In Progress</td>
<td>In Progress</td>
<td>Mastered</td>
</tr>
<tr>
<td><strong>Purchases on credit</strong></td>
<td>Not Attempted</td>
<td>Mastered</td>
<td>Mastered</td>
</tr>
<tr>
<td><strong>Purchases using cash and credit</strong></td>
<td>In Progress</td>
<td>Mastered</td>
<td>Mastered</td>
</tr>
</tbody>
</table>
Assessment can be shown to student
(Carnegie Learning’s Cognitive Algebra I Tutor)
Outline

• Introduction to ITS
• Why are ITS effective?
  – Feedback and hints on steps (not just answers)
  – Adaptive task selection
Why are ITS uncommon?
  – Model development is costly
  – ITS disrupt classroom social structure
• Can crowd-sourcing help solve ITS’s problems?
How does an ITS know which steps to expect?

- Tool (e.g., editor) emits student steps
- Expected steps (correct + misconceived)
- Student competences, affect, etc
- Hint & feedback generator
- Comparison
- Next task selector
Sources of expected steps

• Model-tracing
  – Human author represents each problem formally
  – & develops knowledge-based (e.g., rules) model that can solve every problem in all possible ways

• Constraint-based tutoring
  – Like model-tracing, but human author provides a single solution instead of a formalized problem

• Example-tracing (scripting)
  – Human author provides all possible solutions
  – One such “example” per problem
Problems with each source

• Model-tracing & constraint-based tutoring
  – Very expensive to develop the knowledge-base
• Example tracing
  – Quality control
  – Maintenance
  – Can this workshop solve these problems?
Outline

• Introduction to ITS
• Why are ITS effective?
  – Feedback and hints on steps (not just answers)
  – Adaptive task selection
• Why are ITS uncommon?
  – Model development is costly
  – ITS disrupt classroom social structure
• Can crowd-sourcing help solve ITS’s problems?
ITS disrupt classroom social structures

• Working together? Individual work is divorced from small-group and whole-class work
  – Collaboration & discussion is difficult when different student are in different units
  – Competition is inevitable with self-paced

• Who’s the boss? ITS teaches tells students what to do but instructor can’t tell ITS what to do
  – Instructors may not like ITS’s steps
  – Instructors may not like ITS’s hints, instruction
  – Instructors are responsible for ITS’s assessment
Can crowd-sourcing help socialize ITS?

- Getting students to work together by having them author ITS examples for each other
  - ITS + forum + ???
  - Learning by Authoring ITS (LAITS) project
- Making teachers the boss of the ITS
  - Teachers provide the ITS examples
  - ITS tells teachers the feedback & hints
  - Formative Assessment with Computer Technology (FACT) project
• Crowd-sourcing may overcome ITS problems
  – Cheaper to provide examples than model
  – Students & teachers provide the examples, get ownership and collaboration
• And the ITS should still be effective
  – Still gives feedback & hints on steps
  – Still can do dynamic task selection based on embedded assessments (maybe)