

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

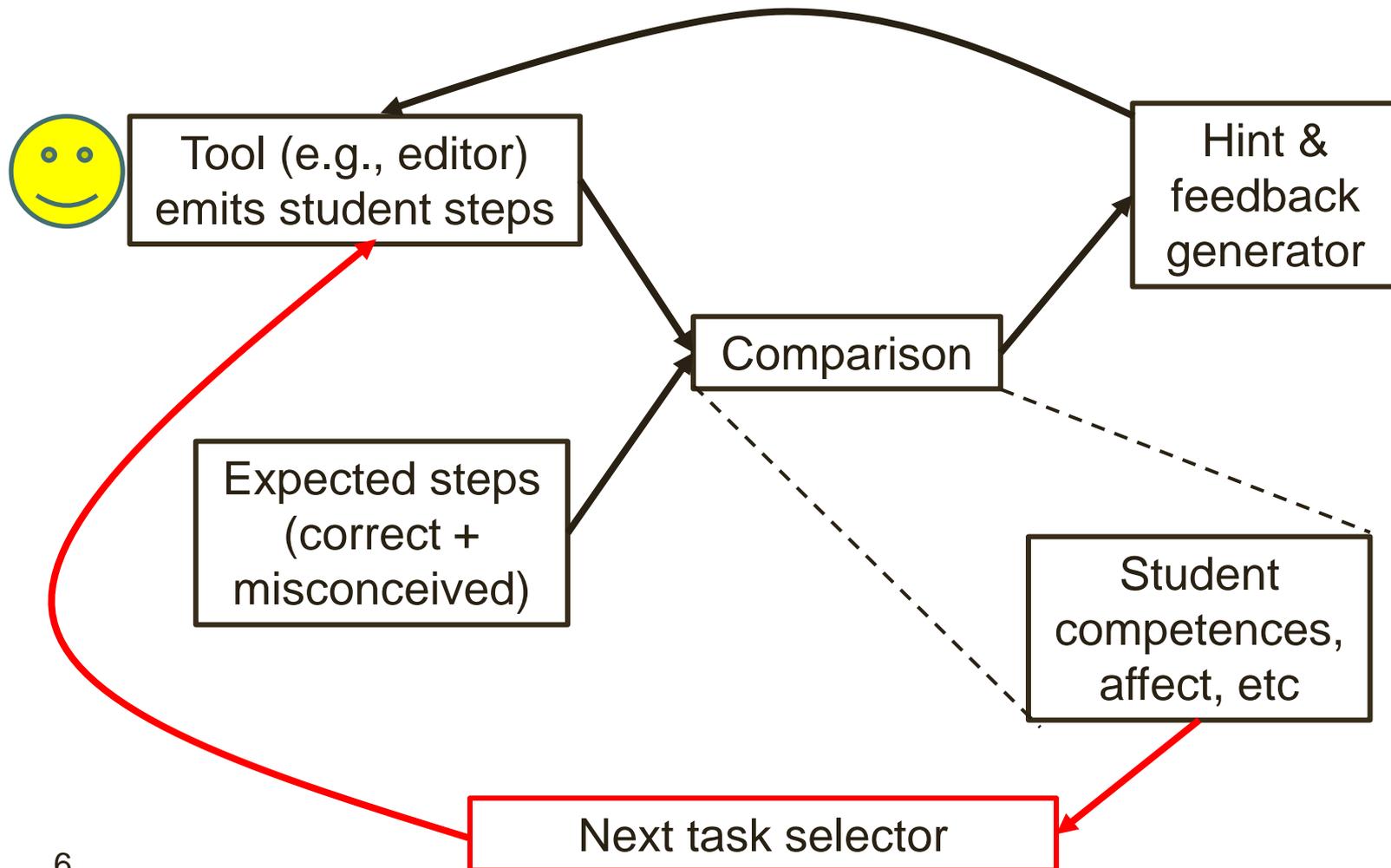
Adding an ITS makes these changes:

- A sequence of modules, each comprised of a **sequence** of
 - Passive media (text, multimedia, lecture)
 - **Exercises** complex, dynamically scaffolded
 - Discussion
 - **Quiz** unnecessary
- 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
 - Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503-524). Charlotte, NC: Information Age Publishers
- 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?



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

Problem

Change Database | New Problem | History | Student Model | Run Query | Help | Log Out

Problem 30 List the titles and numbers of all movies that have won at least one Academy Award and have been made in or after 1988.

SELECT title, number

FROM movies

WHERE aawon>1 and year>=1988

GROUP BY

HAVING

ORDER BY

Feedback Level: Hint

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?

Step

Step

Step

Submit!

Feedback

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*.

The database that the problem refers to

Table Name	Attribute List
<u>DIRECTOR</u>	<u>number</u> lname fname born died
<u>MOVIE</u>	<u>number</u> title type aanom aawon year critics <i>director</i>
<u>STAR</u>	lname fname <u>number</u> born died city
<u>CUSTOMER</u>	lname fname <u>number</u> address rentals bonus jdate
<u>TAPE</u>	<u>code</u> <i>movie</i> pdate times <i>customer</i> hiredate
<u>STARS_IN</u>	<i>movie</i> <u>star</u> role

Cognitive Algebra I Tutor has multiple tools

www.carnegielearning.com

Problem

A rock climber is currently on the side of a cliff 67 feet off the ground. She can climb on average about two and one-half feet per minute.

- 1 When will she be 92 feet off the ground?
- 2 In twenty minutes, how many feet above the ground will she be?
- 3 In 75 seconds, how far above the ground will she be?
- 4 Ten minutes ago, how far above the ground would she

Step: Label a column

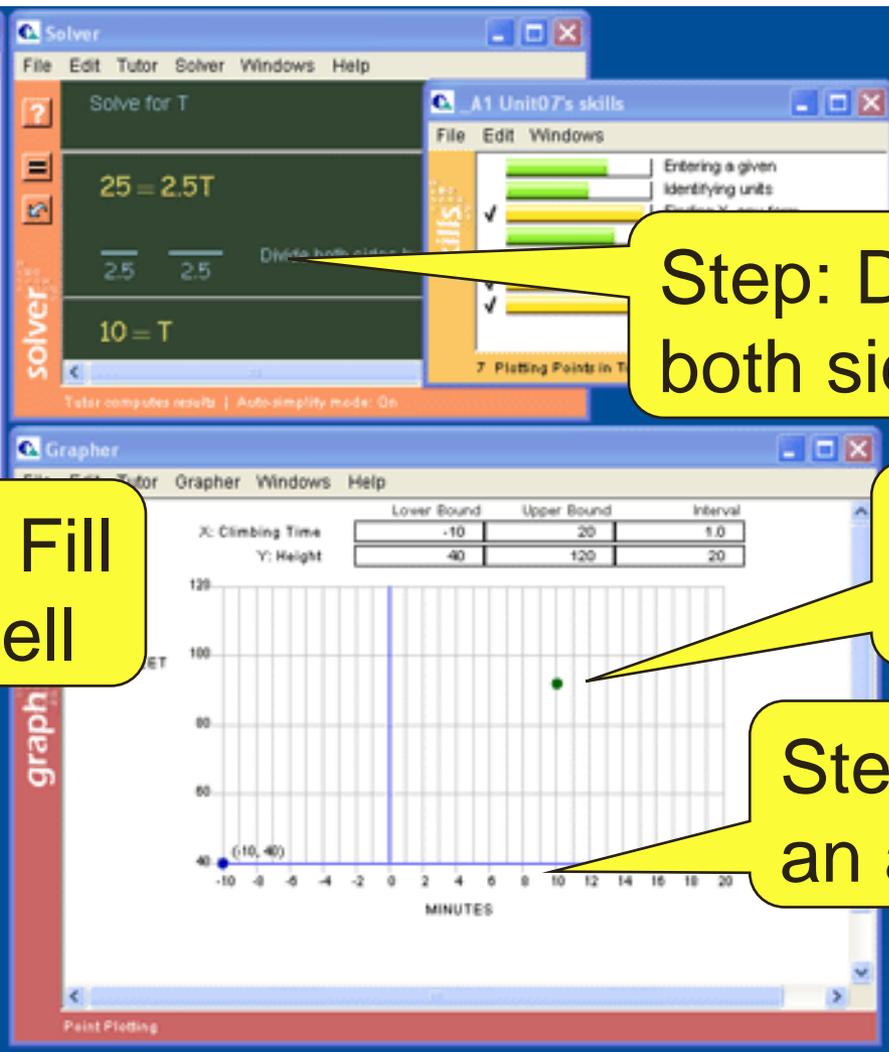
Step: Fill in a cell

Step: Divide both sides

Step: Plot a point

Step: Define an axis

Quantity Name	CLIMBING TIME	HEIGHT OFF GROUND
Unit	MINUTES	FEET
Expression	T	$67 + 2.5T$
Question 1	10	92
Question 2	20	117
Question 3	1.25	70.125
Question 4	-10	42



Andes3 is like power-point, but gives correct (green) vs. incorrect (red) feedback

s5a

Physics

Submit

A model airplane hangs from two strings S1 and S2 which are attached to the ceiling. String S1 is inclined at 45.0 deg, and string S2 is inclined at 60.0 deg, as shown in the figure below.

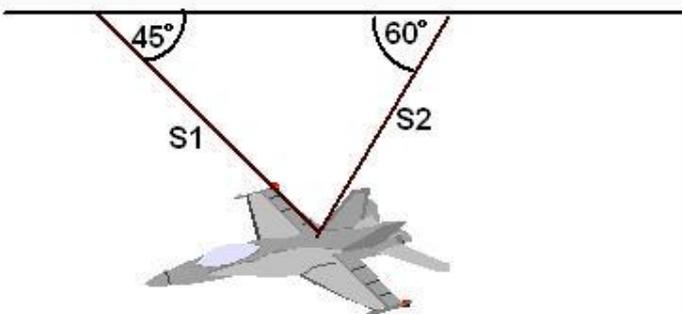
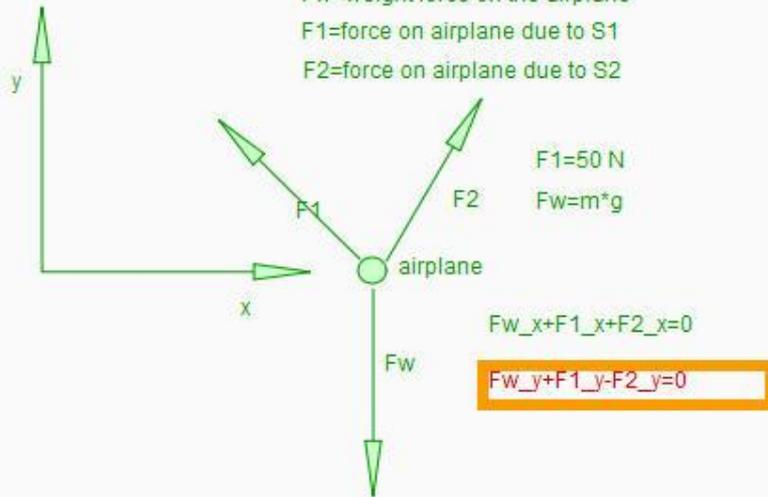
If the tension in string S1 is 50.0 N,

a) find the tension in string S2.

Answer:

b) find the mass of the airplane.

Answer:



Help

Score: 57%

If you need help, click the help button ? below. Click the button above to hide this window.

Check your signs.

[Explain more](#)

Perhaps the sign of the F_{2_y} term should be changed.

help messages

Click here for "what's wrong" and "next step help"

VanLehn, K., Lynch, C., Schultz, K., Shapiro, J. A., Shelby, R. H., Taylor, L., et al. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence and Education*, 15(3), 147-204.

ITS can tutor steps in an argument/essay

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., et al. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments and Computers*, 36(180-193).

The screenshot shows the AutoTutor interface. On the left is a 3D avatar of a male tutor. The main window contains a task question: "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." Below this is a text input field where the student has typed: "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." On the bottom left, a "Log of previous turns" window shows a history of the conversation. Three yellow callout boxes with black outlines point to specific elements: one points to the tutor's avatar and is labeled "Tutor"; another points to the task question and is labeled "Task"; a third points to the student's typed response and is labeled "Student's essay"; and a fourth points to the log window and is labeled "Tutor-student dialogue".

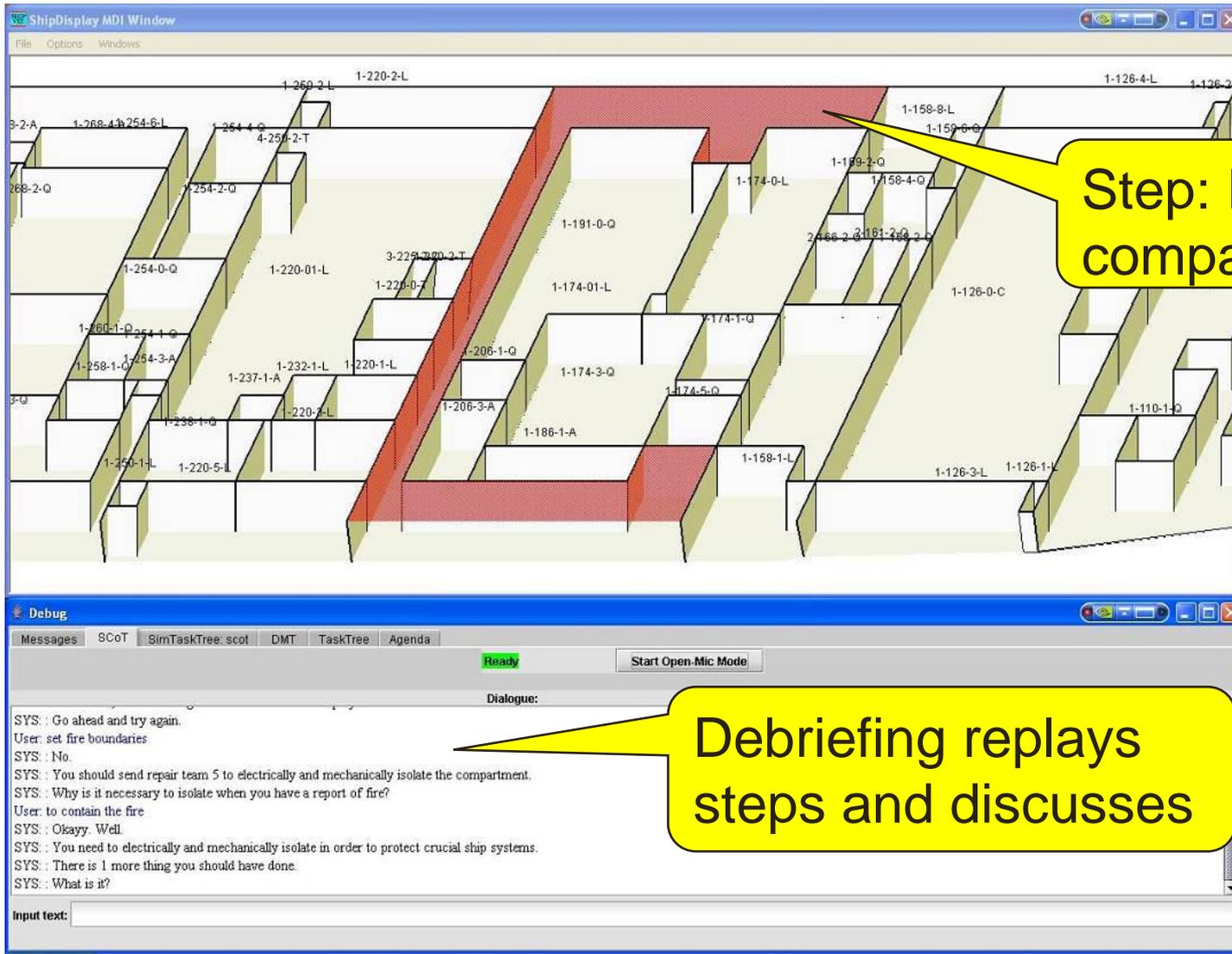
Tutor

Task

Student's essay

Tutor-student dialogue

A simulation-based ITS where all feedback is delayed until done fighting fire



Step: Isolate a compartment

Debriefing replays steps and discusses

Pon-Barry, H., Schultz, K., Bratt, E. O., Clark, B., & Peters, S. (2006). Responding to student uncertainty in spoken tutorial dialogue systems. *International Journal of Artificial Intelligence and Education*, 16, 171-194.

After-action reviews often use a timeline

www.stotler-henke.com

The screenshot displays a 3D terrain simulation. A large red area, representing an Artillery Trajectory, originates from a point labeled "Kaiser Sec SA 2000ft" and extends across the landscape. A yellow callout bubble points to a red diamond marker on the trajectory, stating "ITS marks learning opportunities with red". Below the terrain view is a timeline interface for "Arty1 Conflict: Kaiser Sec intersection with Arty Trajectory". The timeline shows a sequence of events from 17:10:00 to 17:26:00. A red vertical line marks the current time at 17:24:08. The interface includes playback controls, a unit list (Artillery Bty A, Kaiser Sec), and various settings buttons (Details, BSG, RTE, 2D, 2525, Description, Causal Factors, Comments, Reset Cam, Save Cam, Done).

ITS marks learning opportunities with red

ITS can be an non-player character

Nelson, B. C. (2007). Exploring the use of individualized, reflective guidance in an educational multi-user virtual environment. *Journal of Science Education and Technology*, 16(1), 83-97.

The screenshot displays a virtual environment interface for a hospital admissions chart. The interface is divided into several windows:

- 3-D ENVIRONMENT WINDOW:** Shows a 3D scene of a hospital reception desk with a character in a green shirt and a character in a white lab coat.
- TEAM CHAT WINDOW:** Displays chat messages from Dr. Jones and Nurse Patterson.
- INDIVIDUALIZED GUIDANCE SYSTEM:** A window titled "Admissions Chart Help" with a question: "Where are the sick people coming from? Are there any differences in numbers between different areas?"
- Hospital Admissions Chart:** A table with columns: Name, Age, Address, Reason for Visit, and Date.
- TOOLBAR:** A row of icons and navigation buttons.
- HEALTH METER:** A blue bar at the bottom right.

Name	Age	Address	Reason for Visit	Date
Agatha Pearson	54	Tenement #2	diarrhea and stomach pain	10/10/1879
Abby Woods	8	Tenement #2	nightsweats, severe cough, chest pain-- 2nd admission	10/3/1879
Cliff Johnson	51	Tenement #1	Mild stomach pain and	10/12/1879

ITS

ITS

3-D ENVIRONMENT WINDOW

TEAM CHAT WINDOW

INDIVIDUALIZED GUIDANCE SYSTEM

Hospital Admissions Chart

TOOLBAR

HEALTH METER

Learning companion(s) can accompany ITS

Schwartz, D. L., Blair, K. P., Biswas, G., Leelawong, K., & Davis, J. (in press). Animations of thought: Interactivity in the teachable agent paradigm. In R. Lowe & W. Schnotz (Eds.), *Learning with animations: Research and implications for design*. Cambridge, UK: Cambridge University Press.

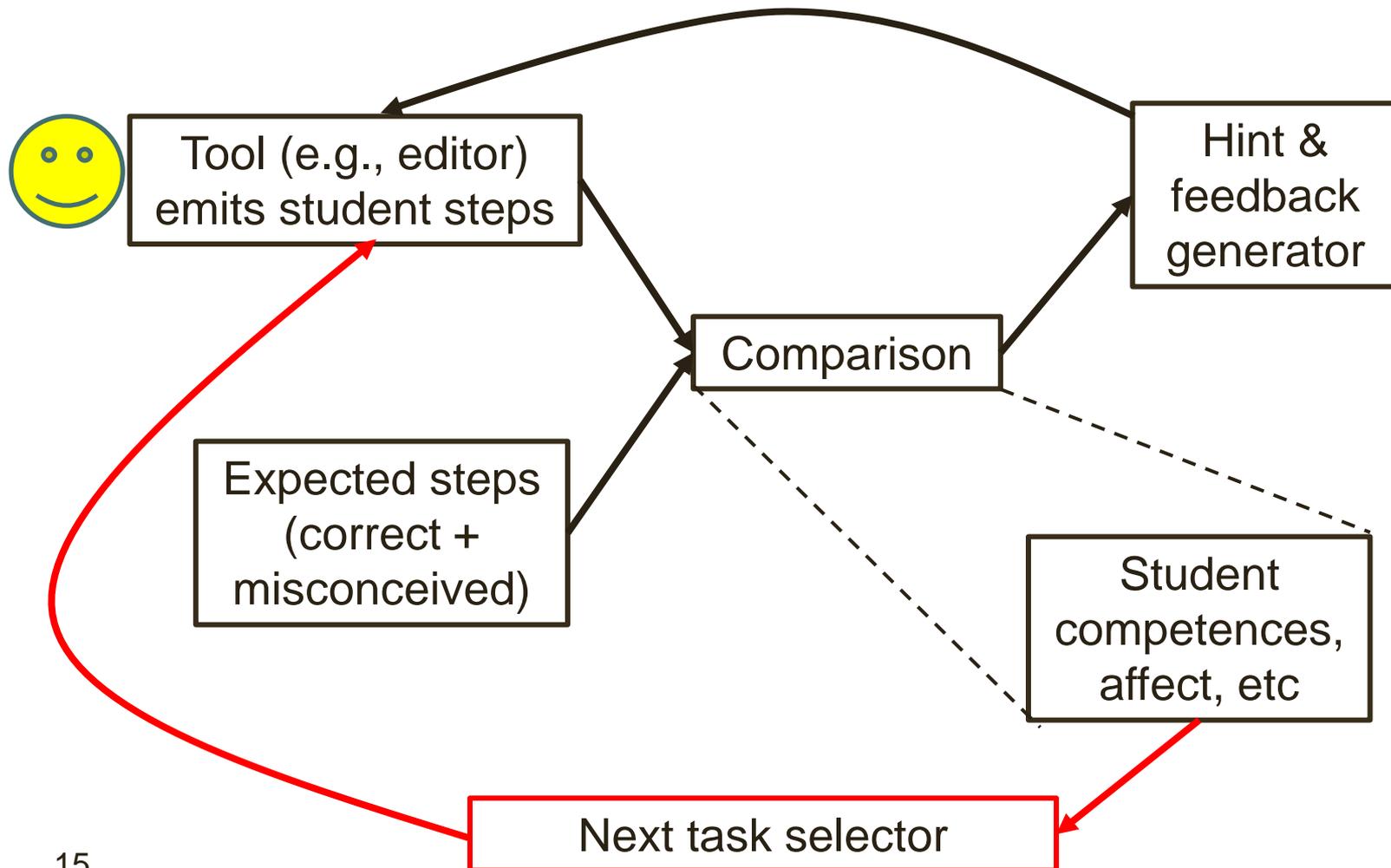
Concept map editor

ITS

Betty, a teachable agent

The screenshot displays the 'Teachable Agents Group at Vanderbilt University' software interface. On the left, a 'Concept map editor' panel includes a toolbar with buttons like 'Porter', 'Teach Concept', 'Teach Link', 'Teach Theme', 'Reverse Link', 'Edit', and 'Delete'. Below the toolbar are 'Themes' for 'artery muscle system', 'cold detection', 'skeletal muscle system', and 'skeletal muscle system'. The main area features a concept map with nodes such as 'cold temperatures', 'cold detection', 'hypothalamus response', 'skeletal muscle activity', 'shivering', 'friction', 'body heat', and 'body temperature', connected by labeled arrows like 'lead to (+)', 'produces (+)', 'increases (+)', 'causes (+)', 'reduces (-)', 'generates (+)', 'decreases (-)', 'preserves (+)', and 'more (+)'. On the left side of the main area, there is a portrait of 'Mr. Davis' (the ITS) and a portrait of 'Betty' (the teachable agent). Below Betty's portrait are buttons for 'Ask', 'Explain', 'Quiz', and 'Repeat'. At the bottom, a 'Talk Log' window shows a conversation: Betty asks 'I should learn more between quizzes. I don't think it helps if I take quizzes too often.' and 'Learning new links helps me understand! How are the other concepts connected?'; the Student asks 'If hypothalamus response increases, what happens to body heat?'; Betty responds 'Based on what I know so far, if hypothalamus response increases, body heat increases a lot. Does my answer make sense?'.

What is the structure of an ITS?



Adding an ITS makes these changes:

- A sequence of modules, each comprised of a **sequence** of
 - Passive media (text, multimedia, lecture)
 - **Exercises** complex, dynamically scaffolded
 - Discussion
 - **Quiz** unnecessary
- 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

Chi, M. T. H. (2009). Active-Constructive-Interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1, 73-105.

Student engagement activity	e.g., history	e.g., algebra equations	Effectiveness
Passive	Listening to a tutor lecture	Watching a tutor's example	Worst
Active	Taking notes on tutor's lecture	Copying an example	OK
Constructive	Answering questions	Solving a problem	Better
Interactive	Discussing questions with a peer or tutor	Solving a problem with a peer or tutor	Best

I > C > A > P

Do computer tutors interact frequently enough, compared to humans?

CAI
(computer
aided
instruction)

- **No** tutoring (baseline for comparisons)
- **Answer**-based tutoring
 - Hints and feedback on short-answer questions

ITS

- **Step**-based tutoring
 - Hints and feedback on steps normally taken when using tool
- **Substep**-based tutoring
 - Tutor can discuss reasoning behind steps
- **Human** tutoring

Answer-based tutoring (CAI)

www.webassign.com

2. + 1/1 points | Previous Answers

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

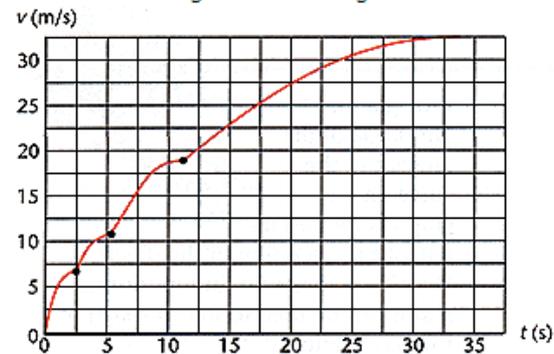


Figure 5-20

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

s5a Physics

A model airplane hangs from two strings S1 and S2 which are attached to the ceiling. String S1 is inclined at 45.0 deg, and string S2 is inclined at 60.0 deg, as shown in the figure below.

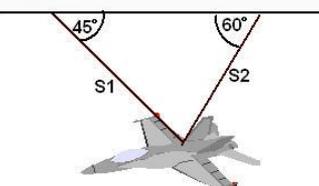
If the tension in string S1 is 50.0 N,

a) find the tension in string S2.

Answer:

b) find the mass of the airplane.

Answer:



Define time T0

g is the acceleration of gravity on earth
 $g = 9.8 \text{ m/s}^2$
 $m = \text{mass of the airplane}$
 $F_w = \text{weight force on the airplane}$

Help Score: 57%

If you need help, click the help button ? below. Click the button above to hide this

Scenario

A rock climber is currently on the side of a cliff 67 feet off the ground. She can climb on average about two and one-half feet per minute.

- When will she be 92 feet off the ground?
- In twenty minutes, how many feet above the ground will she be?
- In 75 seconds, how far above the ground will she be?
- Ten minutes ago, how far above the ground would she have been?

To write the expression, define a variable for the climbing time and use this variable to write a rule for her height above the ground.

Worksheet for Problem BH1T20

Quantity Name	CLIMBING TIME	HEIGHT ABOVE GROUND
Unit	MINUTES	FEET
Expression	T	67 + 2.5T
Question 1	10	92
Question 2	20	117
Question 3	1.25	70.125
Question 4	-10	42

Solver

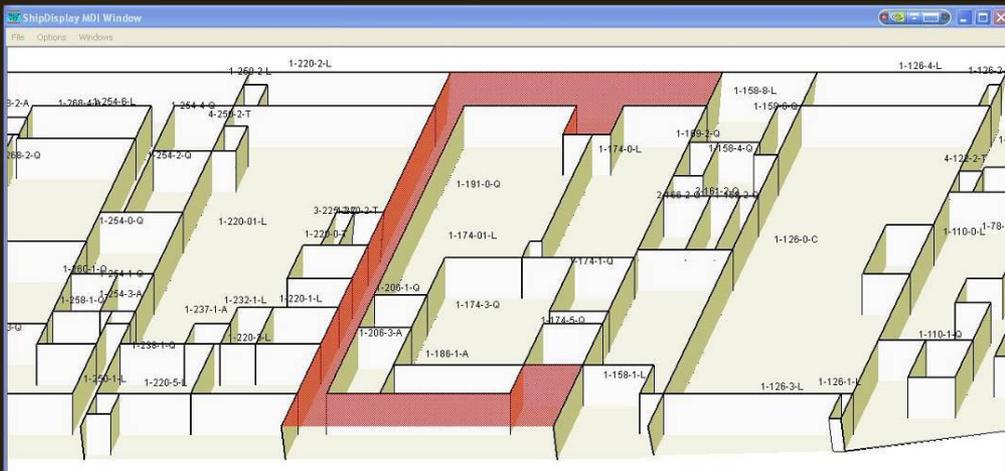
Solve for T

$$25 = 2.5T$$

$$10 = T$$

Grapher

ShipDisplay MDI Window



Debug

Messages SCOT SimTaskTree: scot DMT TaskTree Agenda

Ready Start Open Mic Mode

Dialogue:

SYS: Go ahead and try again.
 User: set fire boundaries
 SYS: No
 SYS: You should send repair team 5 to electrically and mechanically isolate the compartment.
 SYS: Why is it necessary to isolate when you have a report of fire?
 User: to contain the fire
 SYS: Okay. Well.
 SYS: You need to electrically and mechanically isolate in order to protect crucial ship systems.
 SYS: There is 1 more thing you should have done.
 SYS: What is it?

input text:

Substep-based tutoring

VanLehn, K., Jordan, P., & Litman, D. (2007). *Developing pedagogically effective tutorial dialogue tactics: Experiments and a testbed*. Paper presented at the SLaTE Workshop on Speech and Language Technology in Education, Farmington, PA.

Student Interface

Problem Statement | Dialog History | Variables | Equations

You said: definition of kinetic energy

Tutor said: Okay. Please write the equation for how the definition of kinetic energy applies to this problem at T1.

You said: value='ke1=1/2*m*v1^2'

Tutor said: Now it is easy to calculate the magnitude of v1. The magnitude of v1 is 5.0 m/s.

Tutor said: Before going on to the next step, let's think about the application of this equation.

Tutor said: Can we infer the direction of the velocity of the rock at T1 from the rock's kinetic energy at T1?

You said: no, we cannot

Tutor said: Excellent! Please explain why.

Only the magnitude of the velocity and not the direction of it is part of the definition of kinetic energy

Problem Statement | Dialog History | Variables

A 0.6 kg rock in space has a velocity of 7.50 m/s at point B. What is the net work done on the rock at point B?

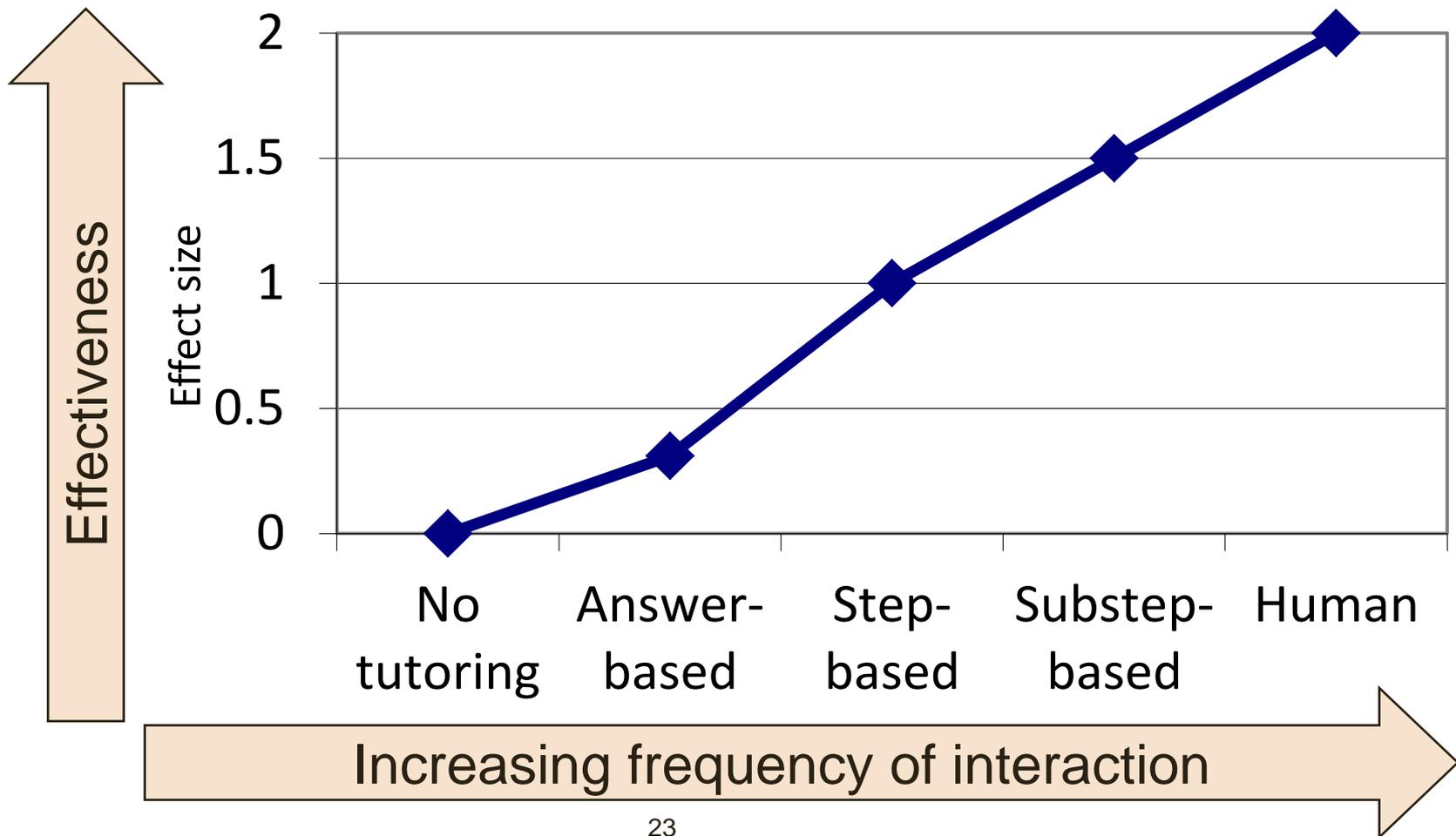
Name | Description

m | The mass of the rock is 0.60 kg

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



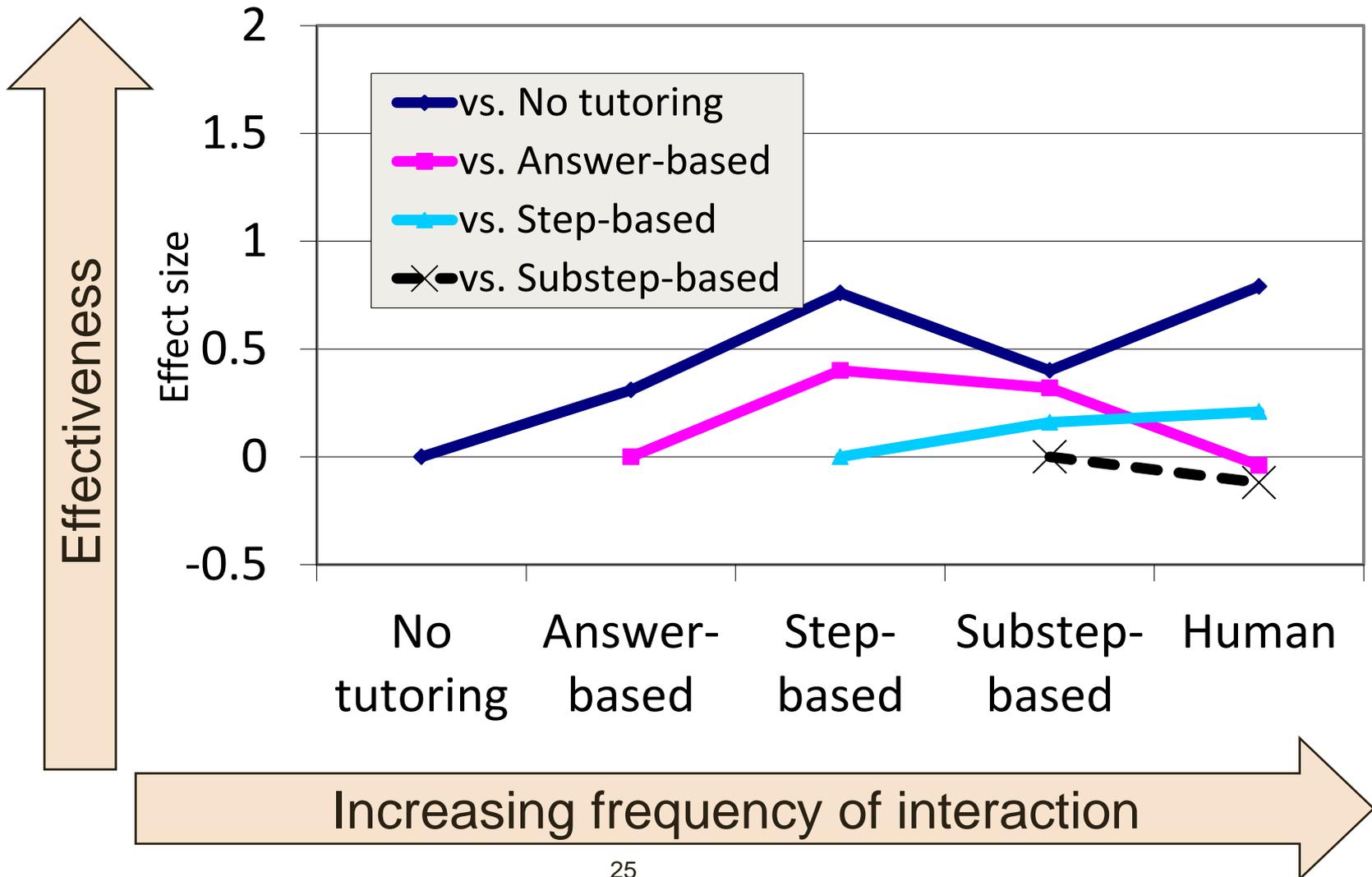
All possible pairwise comparisons

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems. *Educational Psychologist*, 46(4), 197-221.

Tutoring type	vs. other tutoring type	Num. of effects	Mean effect	% reliable
Answer-based	no tutoring	165	0.31	40%
Step-based		28	0.76	68%
Substep-based		26	0.40	54%
Human		10	0.79	80%
Step-based	answer-based	2	0.40	50%
Substep-based		6	0.32	33%
Human		1	-0.04	0%
Substep-based	step-based	11	0.16	0%
Human		10	0.21	30%
Human	sub-step based	5	-0.12	0%

Graphing all 10 comparisons:

none < answer < step = substep = human

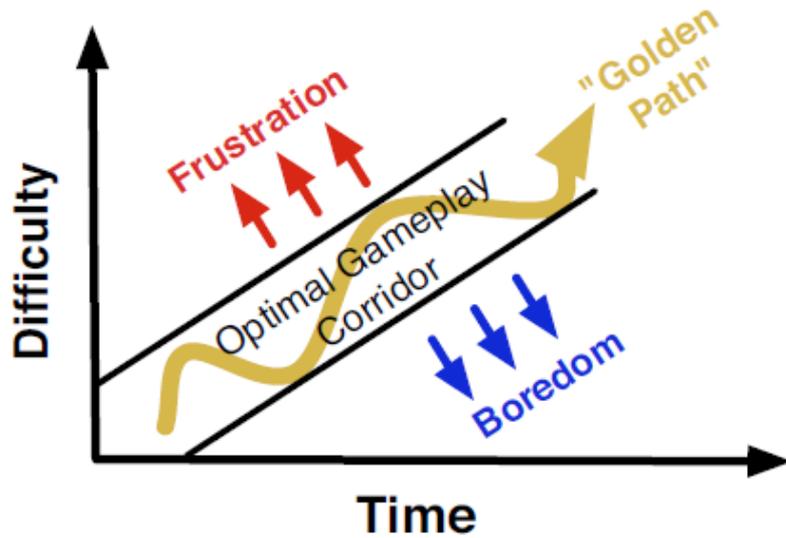


Outline

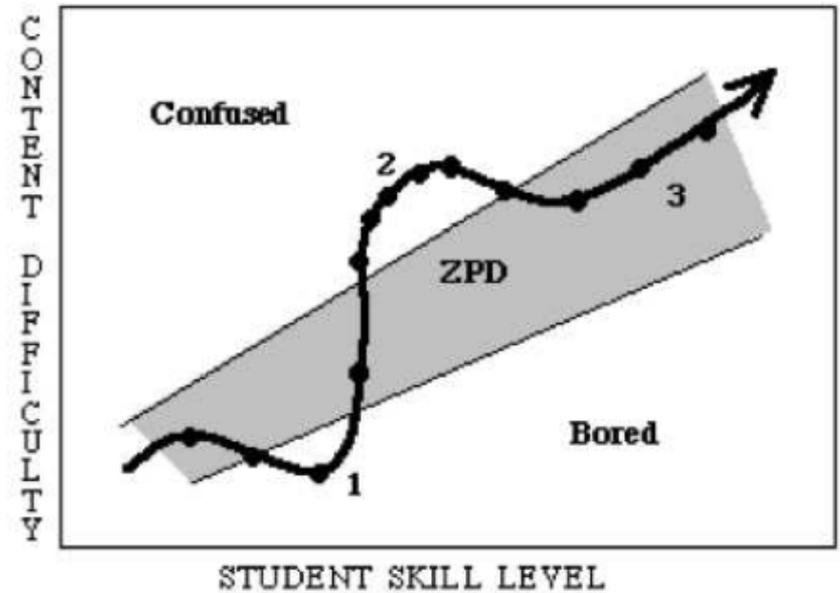
- Introduction to ITS
- Why are ITS effective?
 - Feedback and hints on steps (not just answers)
- Adaptive task selection
- Why are ITS uncommon?
 - Model development is costly
 - ITS disrupt classroom social structure
- Can crowd-sourcing help solve ITS's problems?



The Zone of Proximal Development (ZPD)



(a) Optimal Game Play Corridor

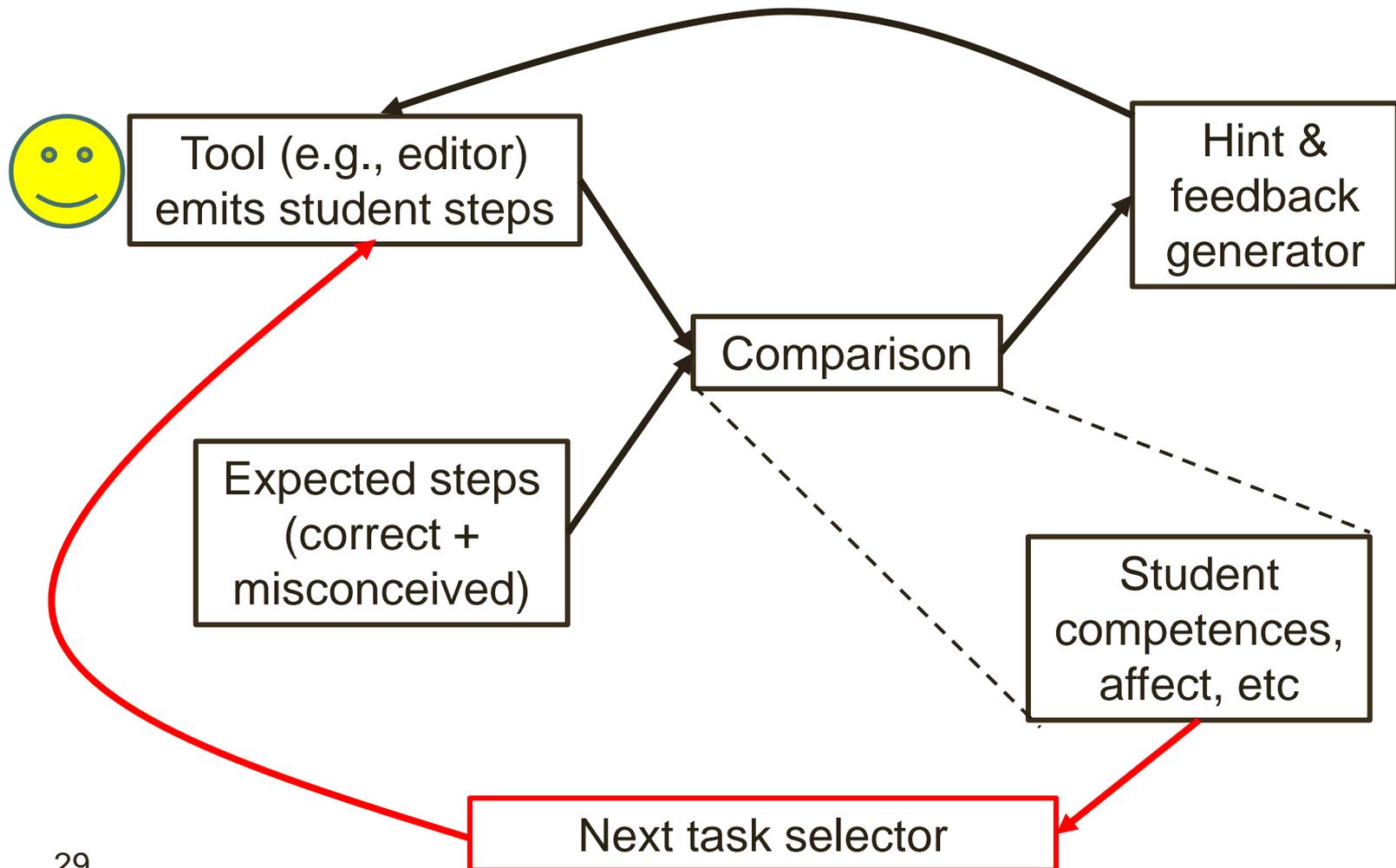


(b) Zone of Proximal Development

How can we keep students in their ZPDs?

- “Mastery learning” means continuing to work on a module until you have mastered it.
 - Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13, 4-16.
- With an ITS, traditional mastery tests are replaced by the stealth assessment

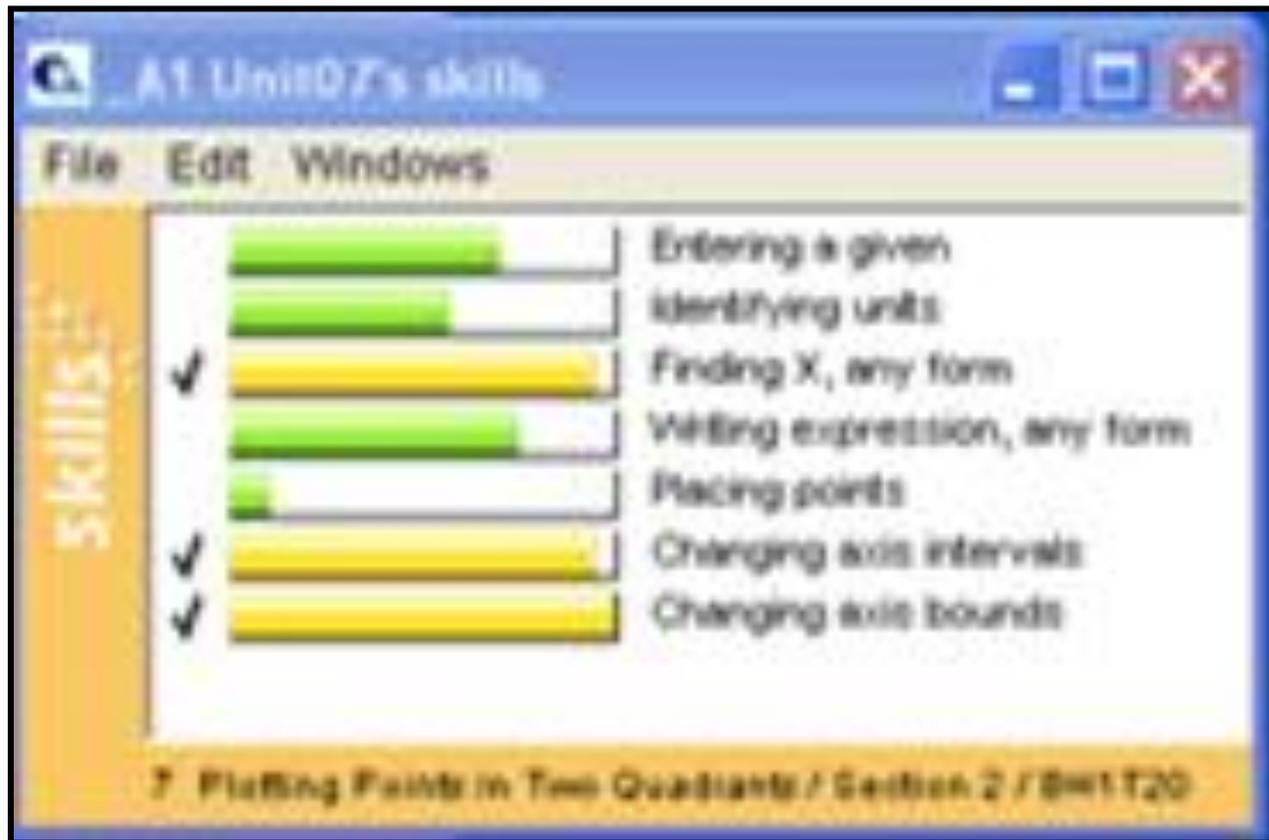
Mastery learning implemented by the red path below



Students may be allowed to work at own pace (Quantum's accounting tutor)

	Jim	Sue	Juan
Borrowing money from bank	In Progress	In Progress	Mastered
Collecting on receivables	Mastered	Mastered	Mastered
Events that aren't transactions	In Progress	In Progress	Mastered
Investments for stock	In Progress	Not Attempted	In Progress
Involving more than two accounts	In Progress	In Progress	Not Attempted
Paying dividends	Mastered	Mastered	Mastered
Paying expenses	Not Attempted	Not Attempted	Not Attempted
Payments to creditors	Mastered	Mastered	Mastered
Prepaying expenses	In Progress	Mastered	In Progress
Providing services to customers	In Progress	In Progress	Mastered
Purchases on credit	Not Attempted	Mastered	Mastered
Purchases using cash and credit	In Progress	Mastered	Mastered

Assessment can be shown to student (Carnegie Learning's Cognitive Algebra I Tutor)



Outline

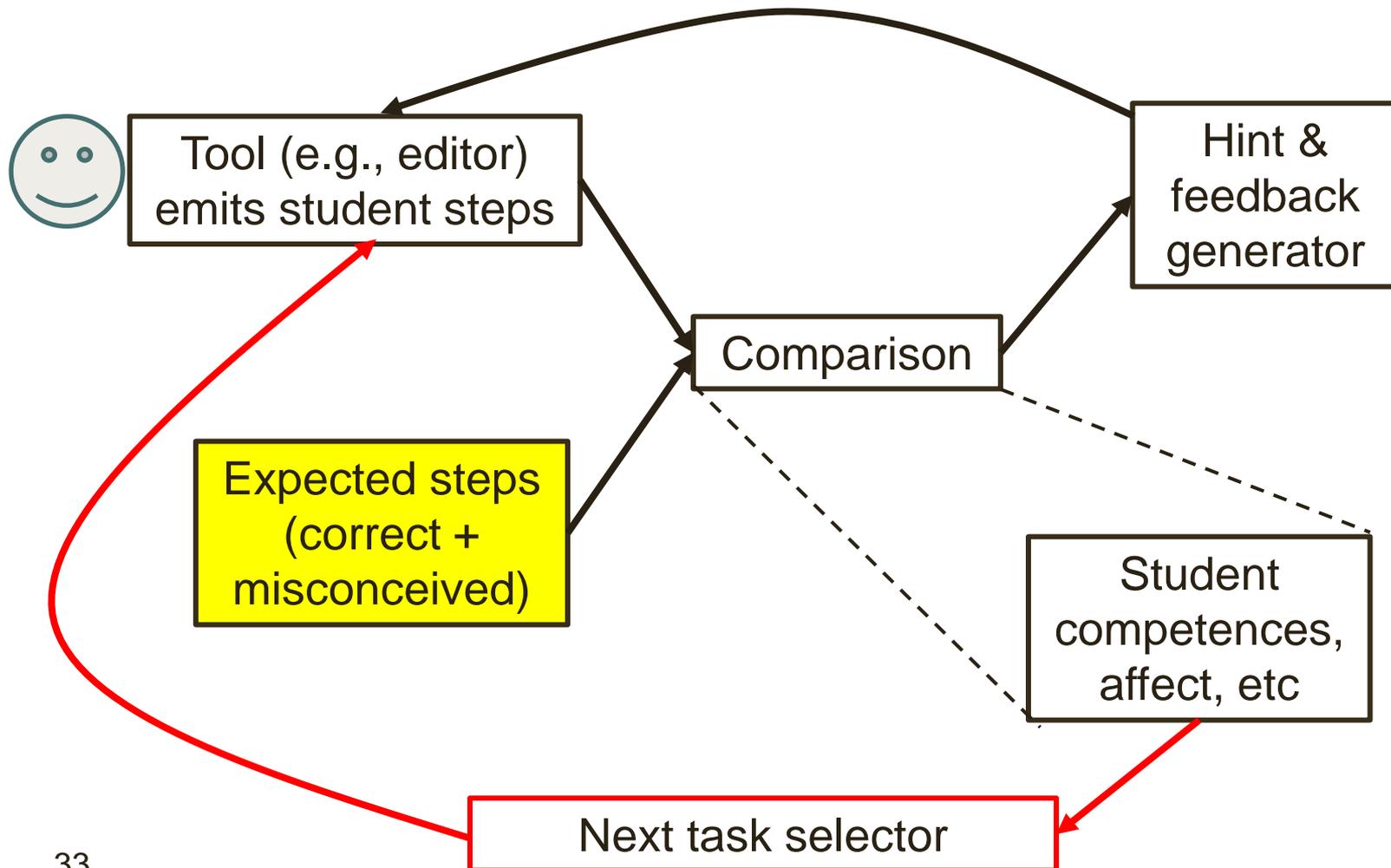
- Introduction to ITS
- Why are ITS effective?
 - Feedback and hints on steps (not just answers)
 - Adaptive task selection



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?



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?



Next

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

Final slide

- 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)