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:

dynamically adaptive

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

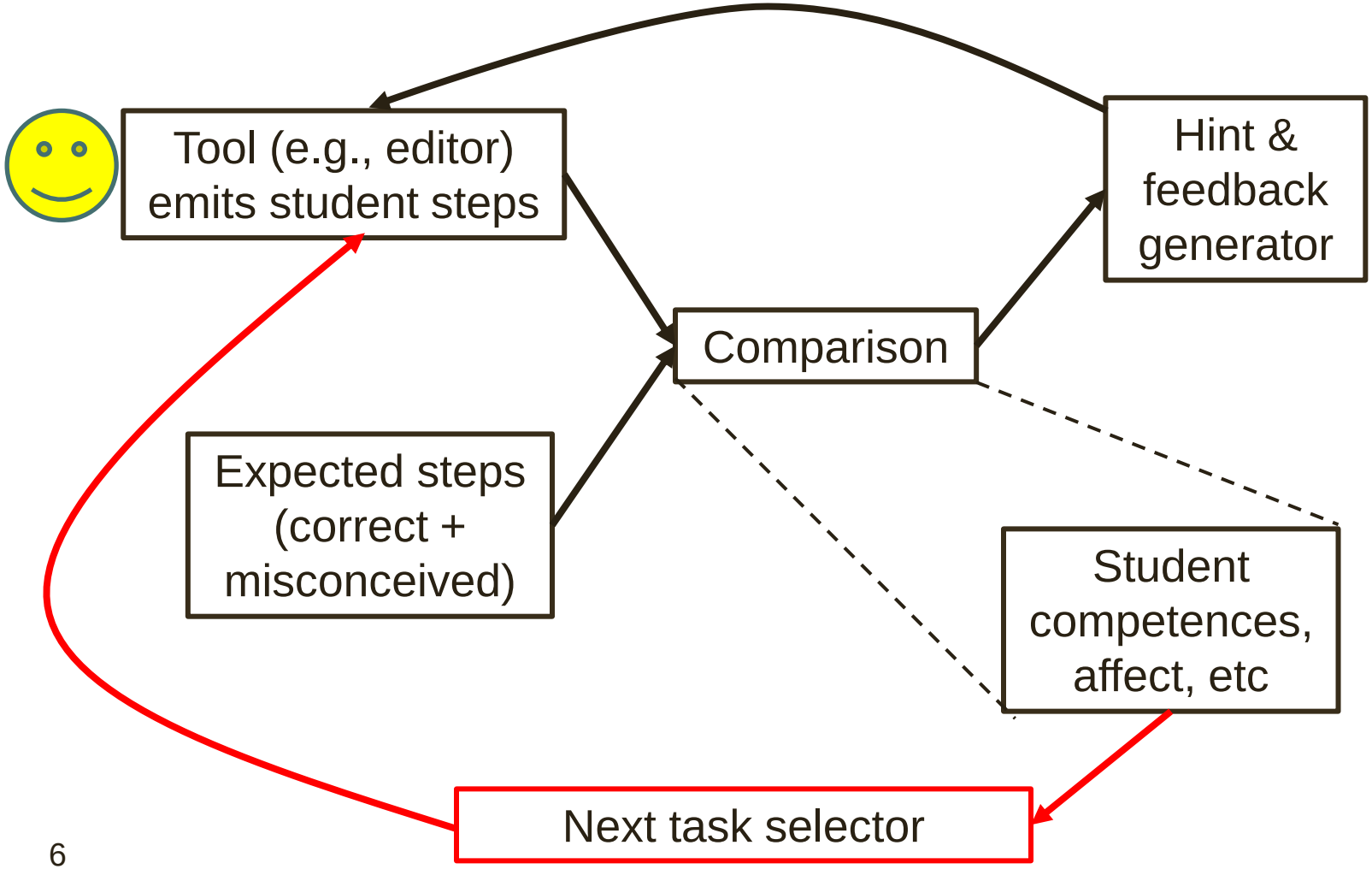
complex, dynamically scaffolded

unnecessary

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 | Submit Answer | Reset

Step

Step

Step

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

Schema for the MOVIES Database

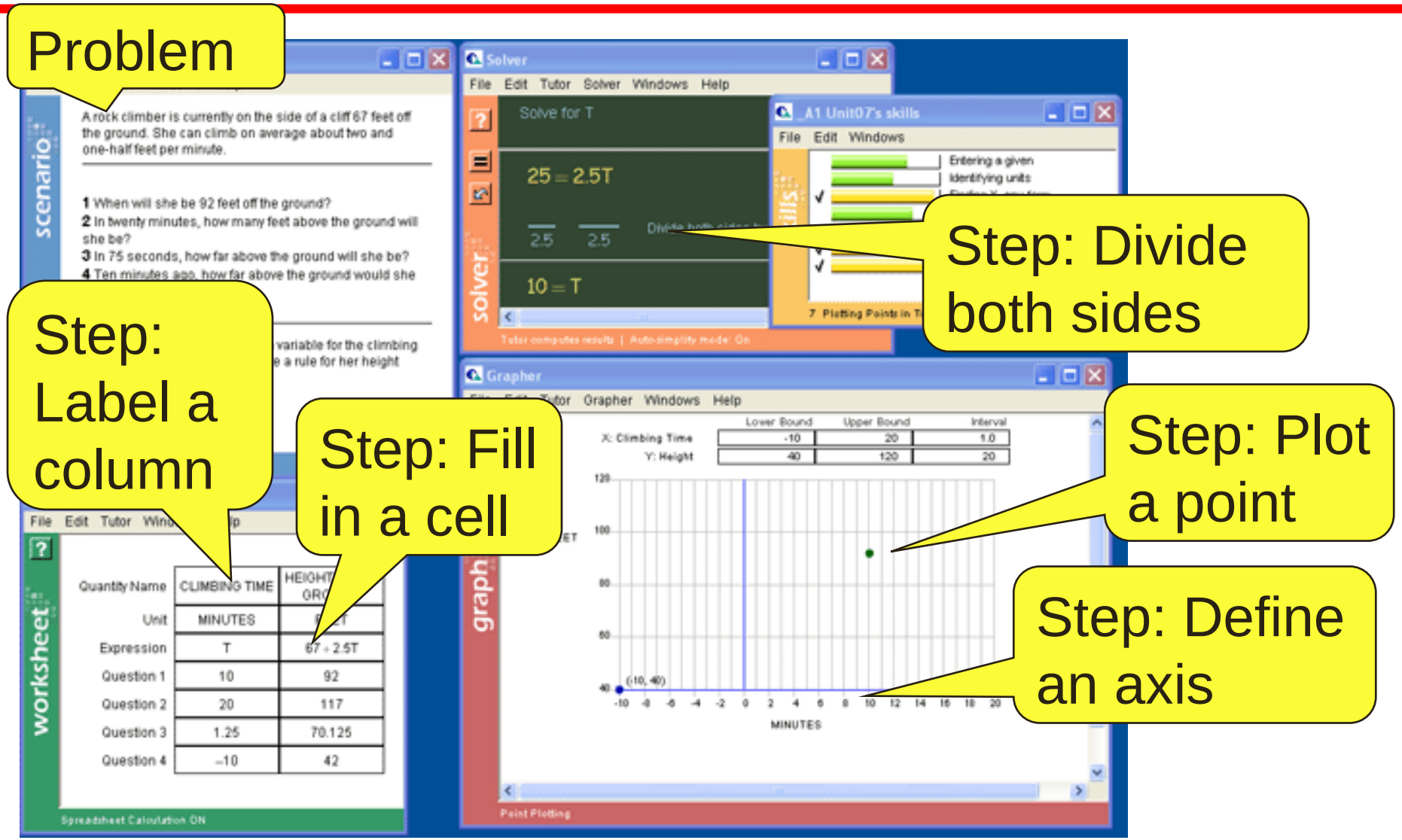
The general description of the database is available **here**. Clicking on the name of a table brings up the table details. Keys in the attribute list are underlined , foreign keys are in *italics*.

Table Name	Attribute List
DIRECTOR	number lname fname born died
MOVIE	number title type aanom 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

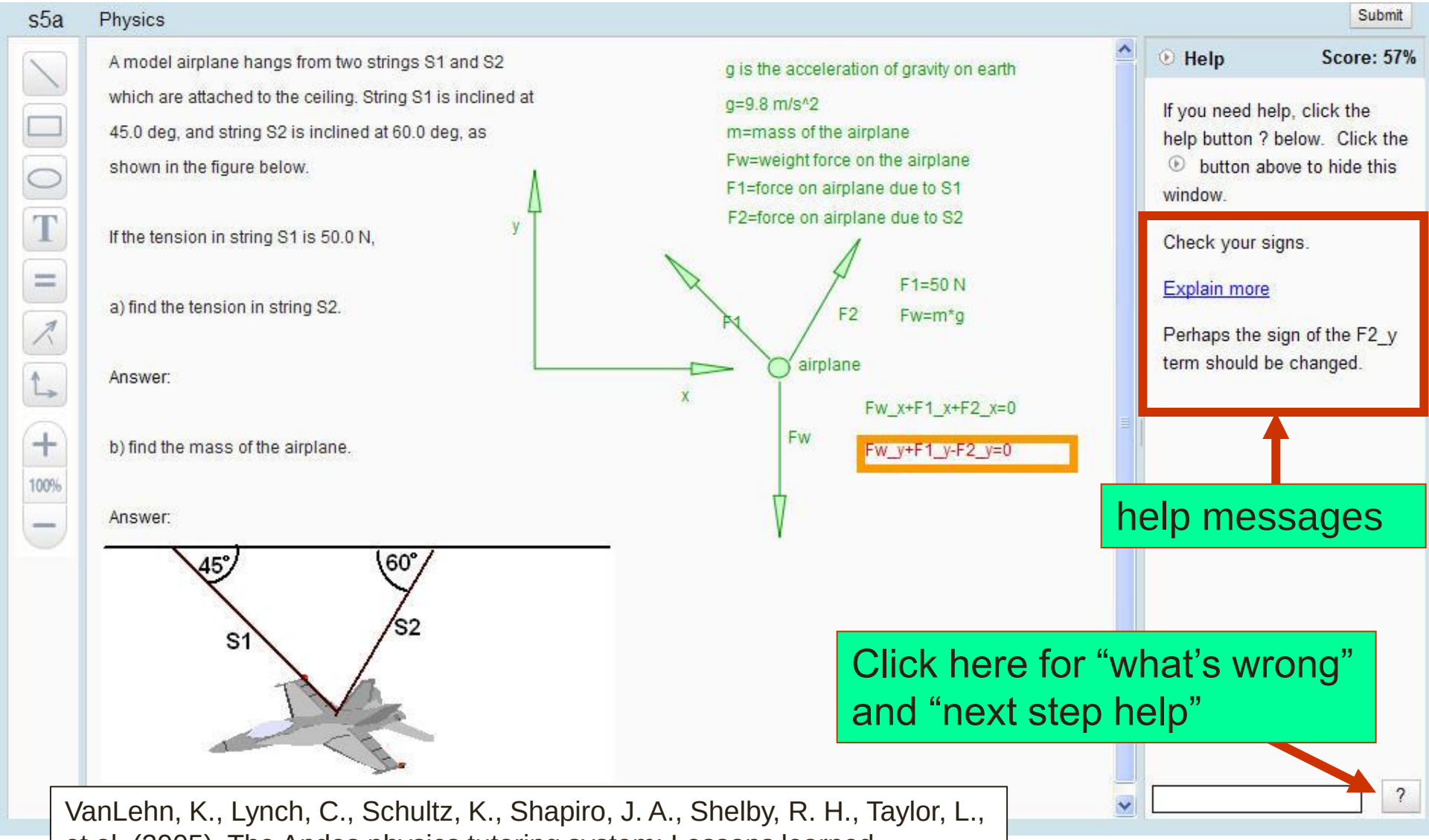
The database that the problem refers to

Cognitive Algebra I Tutor has multiple tools

www.carnegielearning.com



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



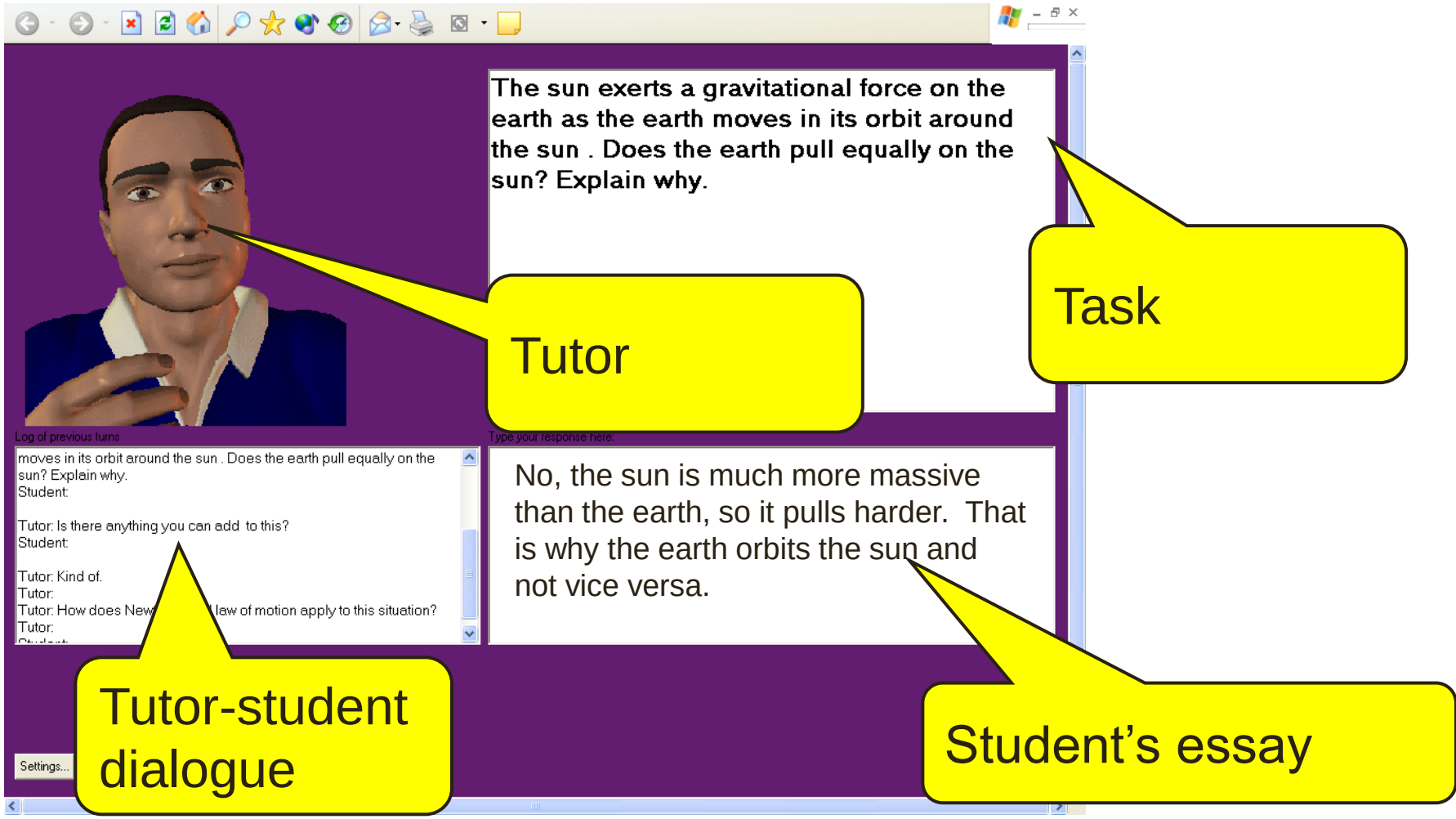
help messages

Click here for “what’s wrong” and “next step help”

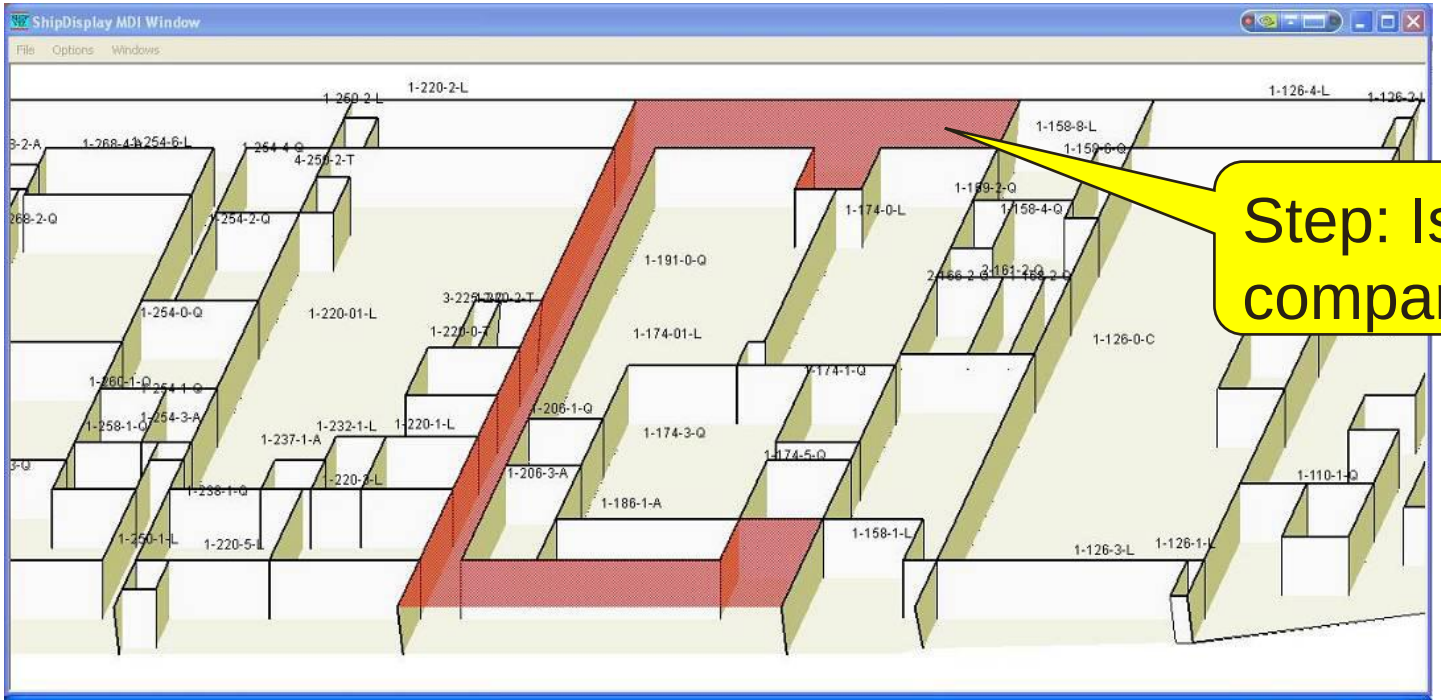
VanLehn, K., Lynch, C., Schultz, K., Shapiro, J. A., Shelby, R. H., Taylor, L., et al. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence and Education, 15(3), 147-204.*

ITS can tutor steps in an argument/essay

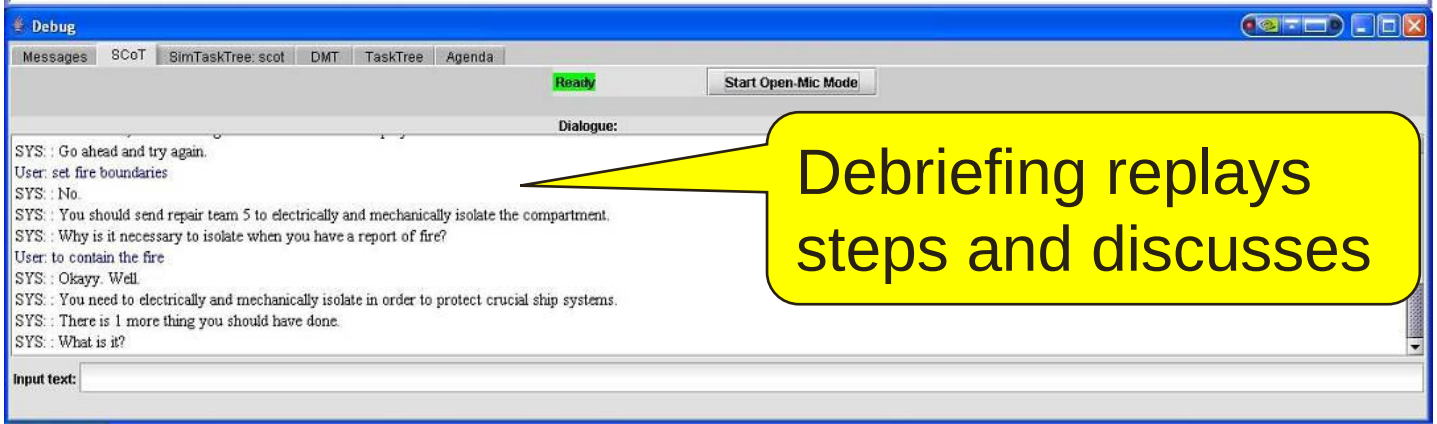
Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., et al. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments and Computers, 36(180-193).*



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



Step: Isolate a compartment

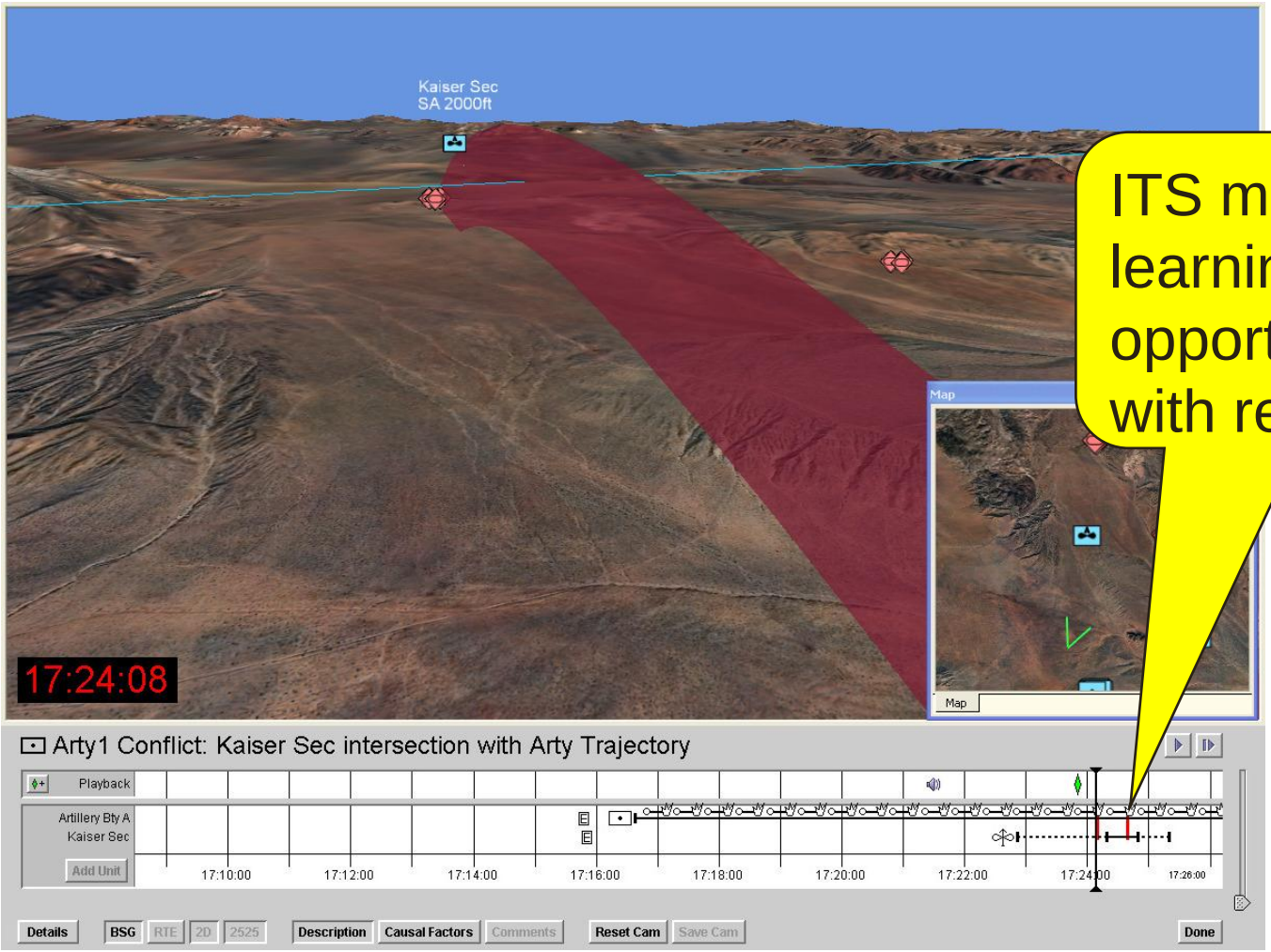


Debriefing replays steps and discusses

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

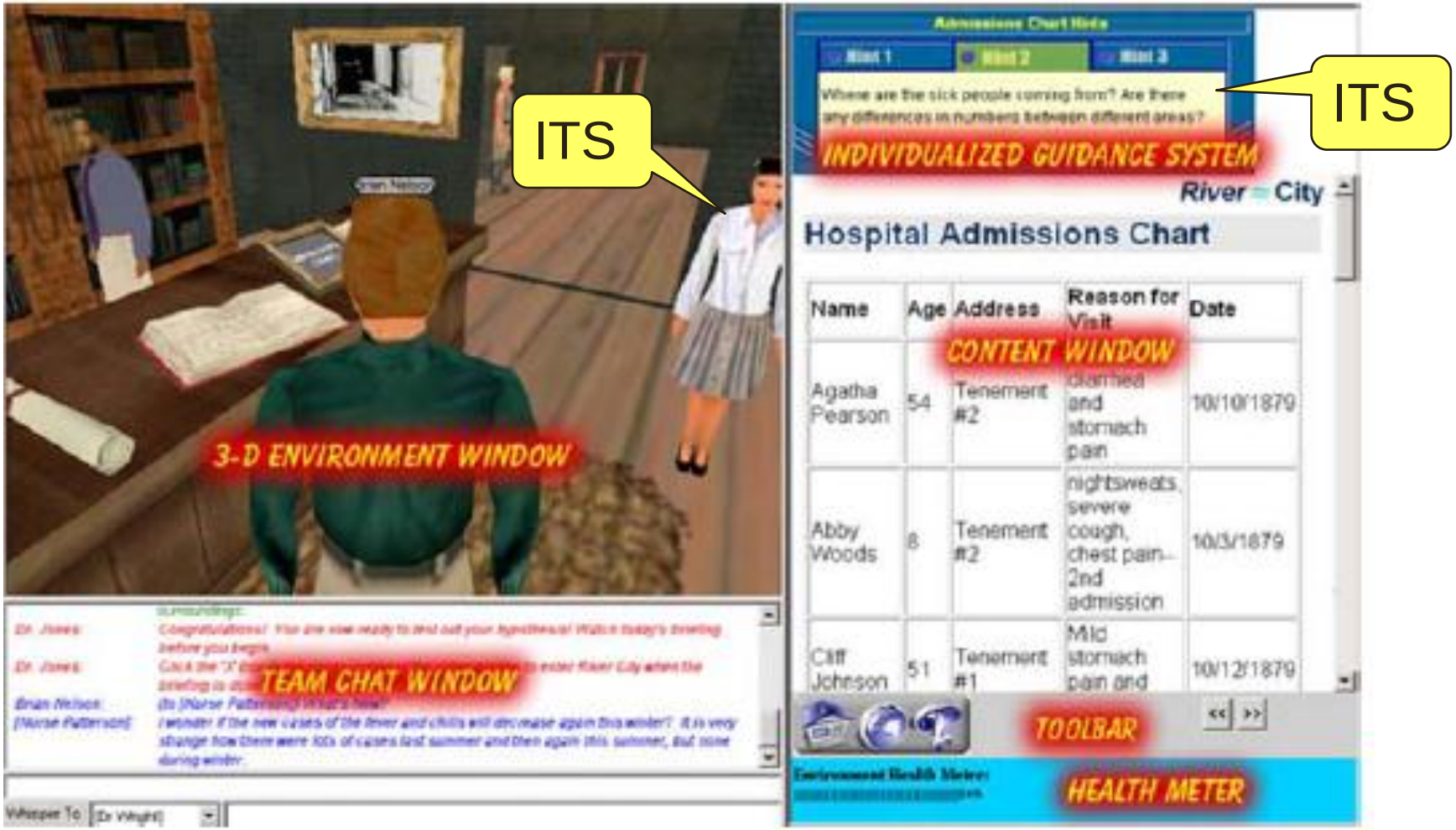
www.stotler-henke.com



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.*



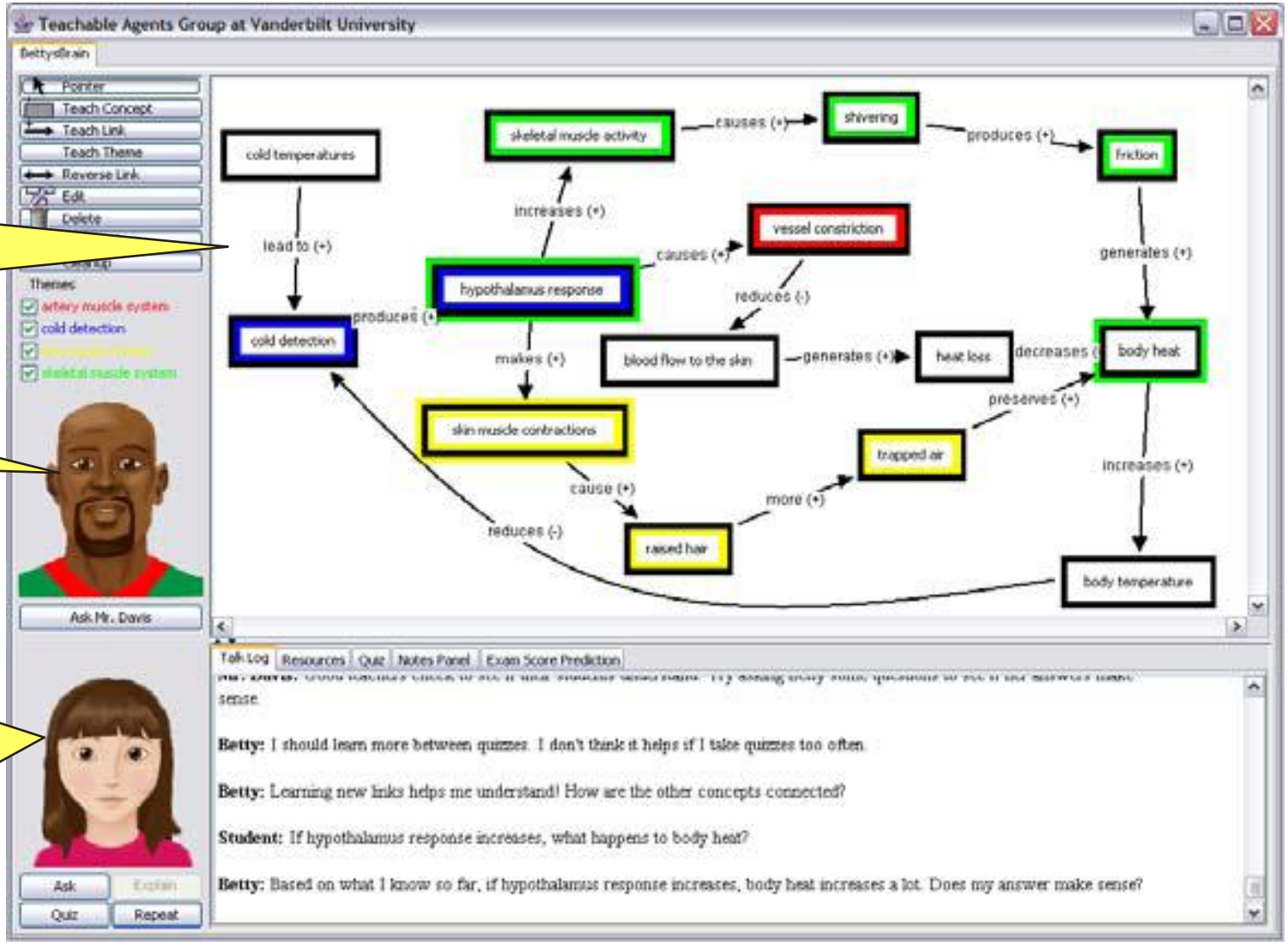
Learning companion(s) can accompany ITS

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.*

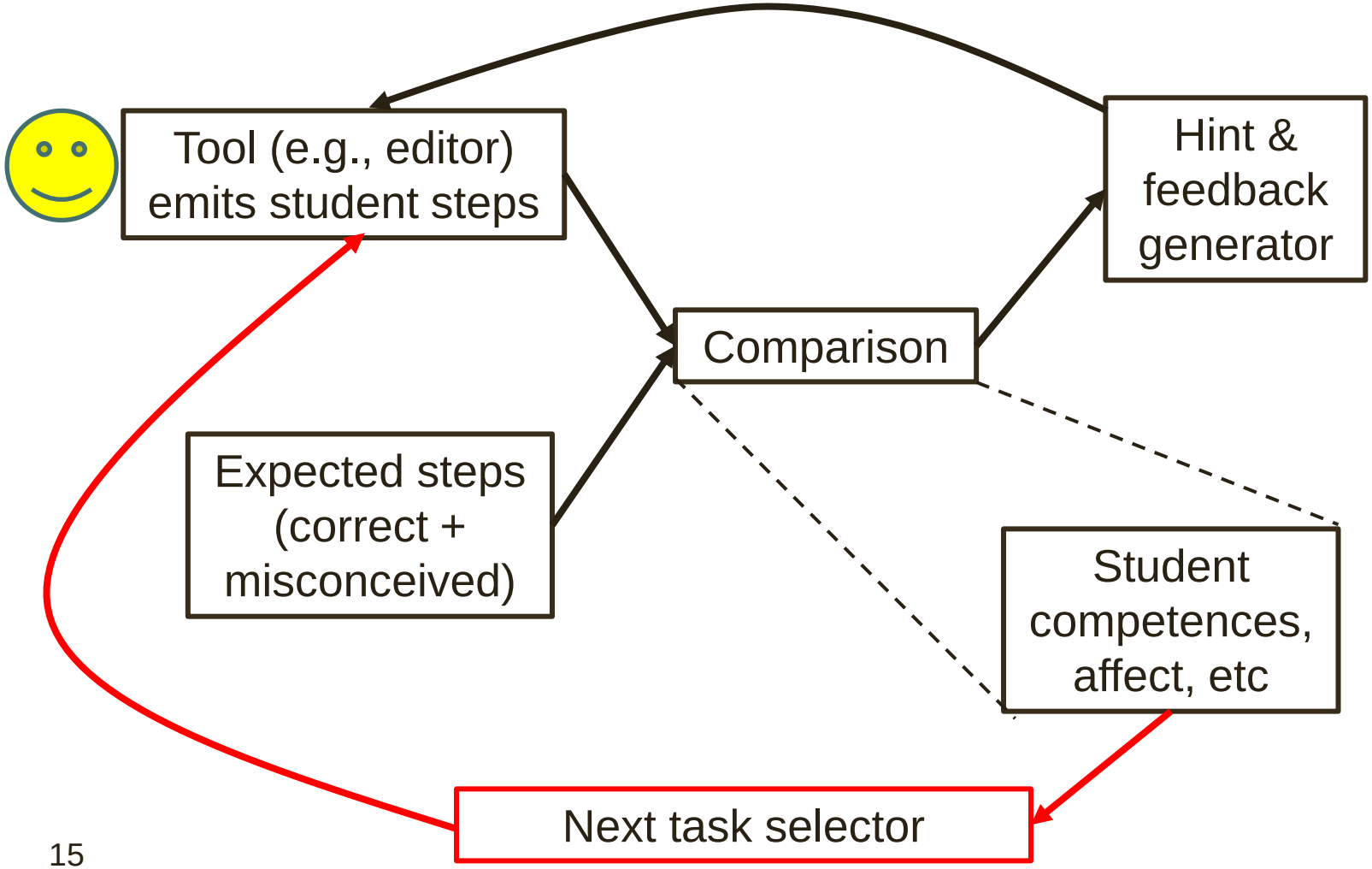
Concept map editor

ITS

Betty, a teachable agent



What is the structure of an ITS?



Adding an ITS makes these changes:

dynamically adaptive

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

complex, dynamically scaffolded

unnecessary

Outline

- Introduction to ITS
- Next 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)

ITS

- **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

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.

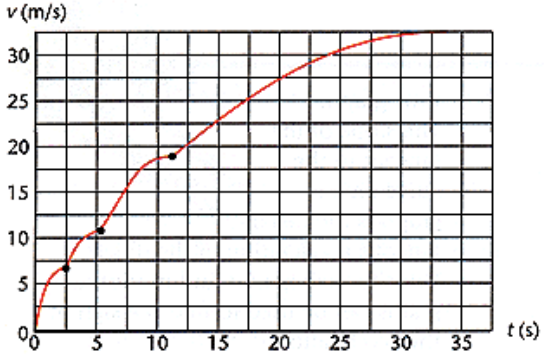
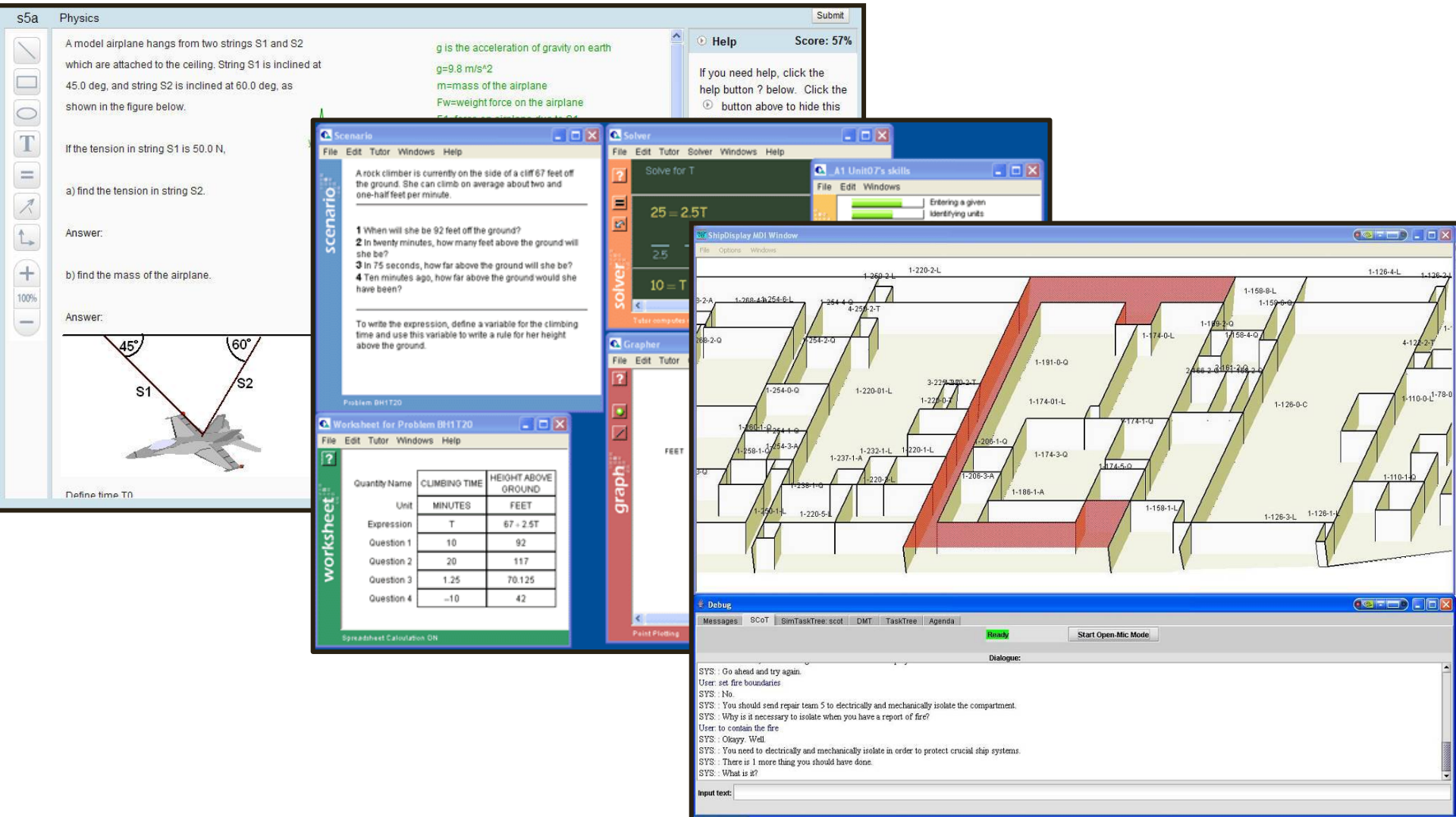


Figure 5-20

Largest acceleration

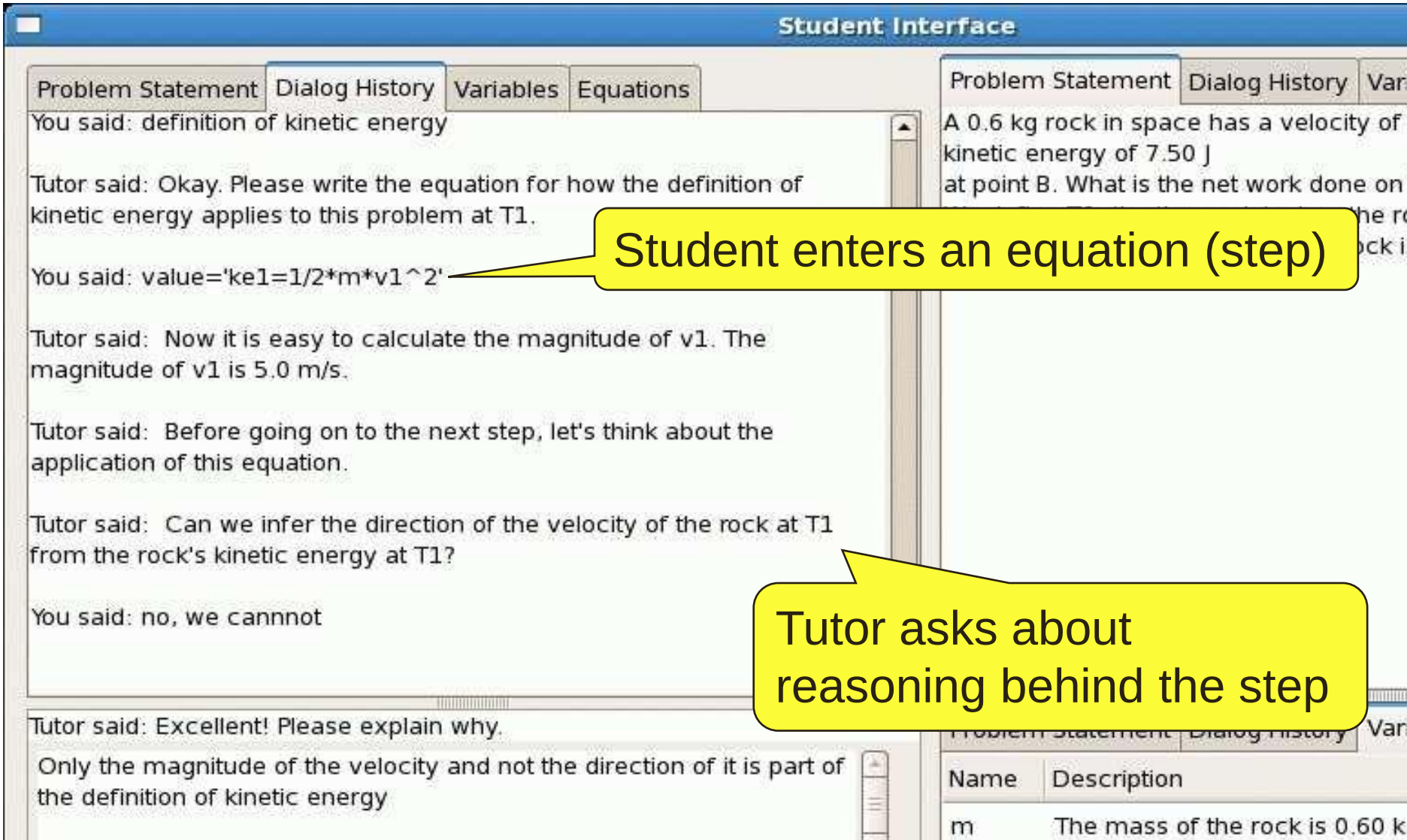
- ○ t = 5.2 s to t = 6.0 s
- ○ t = 10.5 s to t = 11.0 s
- ◉ t = 0 to t = 0.5 s
- ○ t = 31 s to t = 32 s

Step-based tutoring

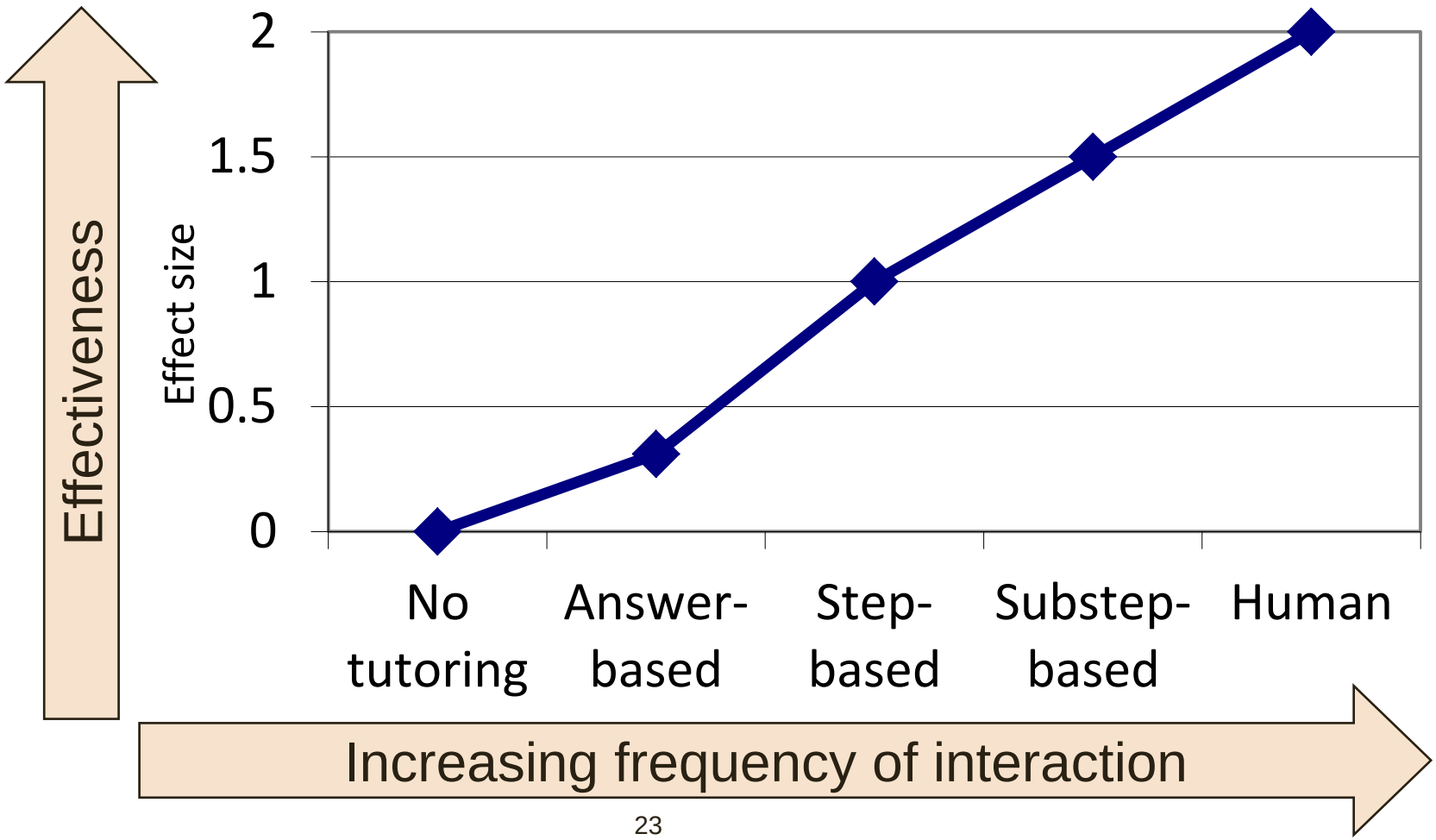


Substep-based tutoring

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.*



A common belief: The more frequent the interaction, the more effective the tutoring

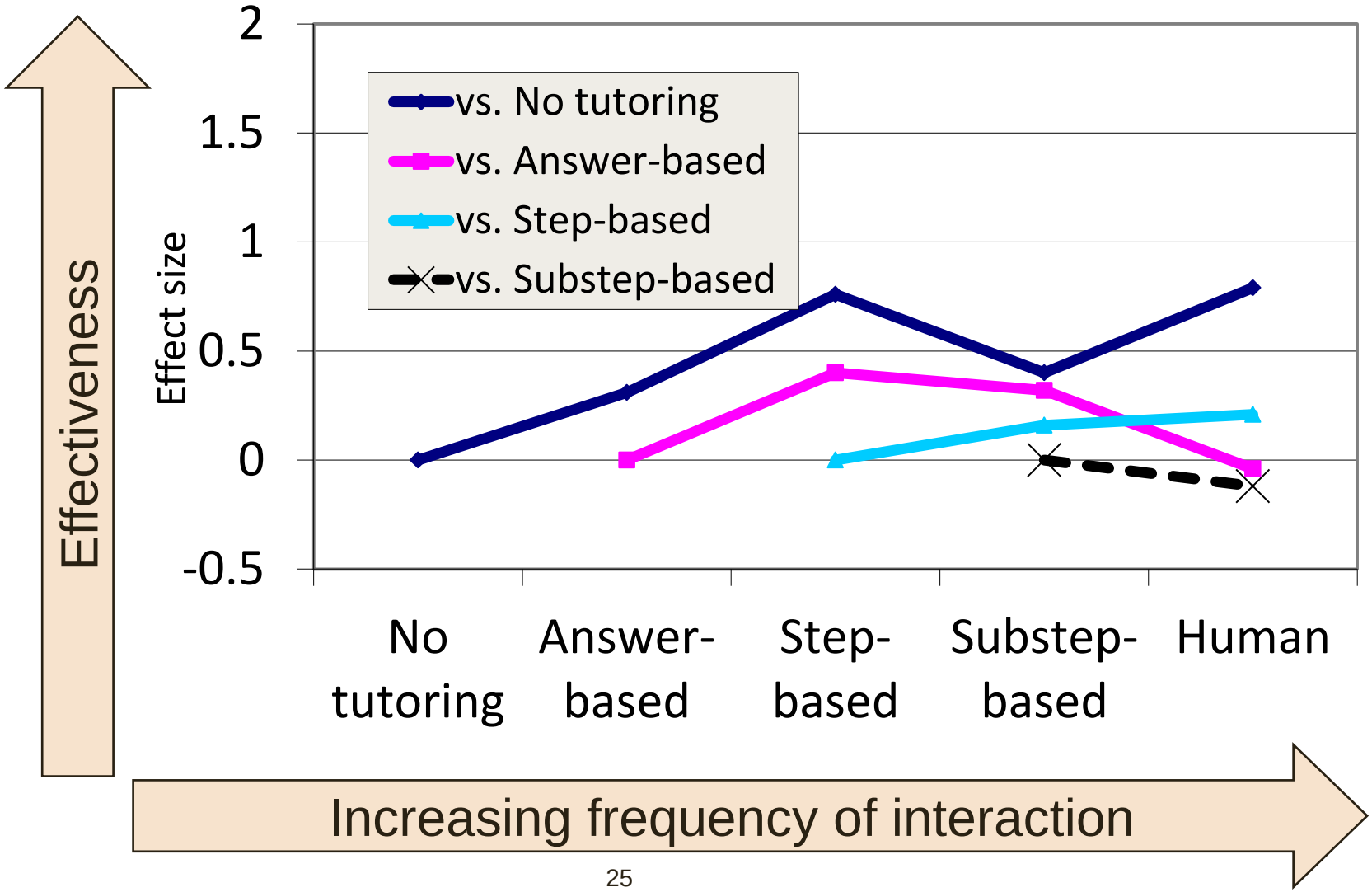


All possible pairwise comparisons

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

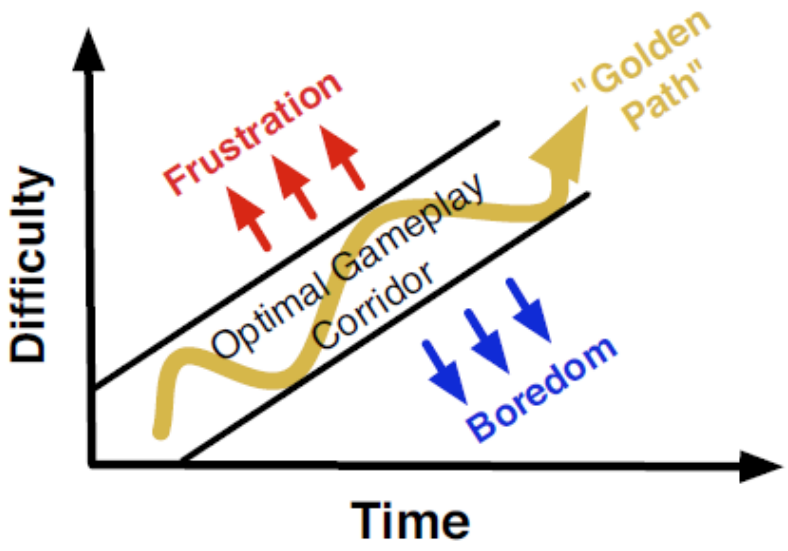


Outline

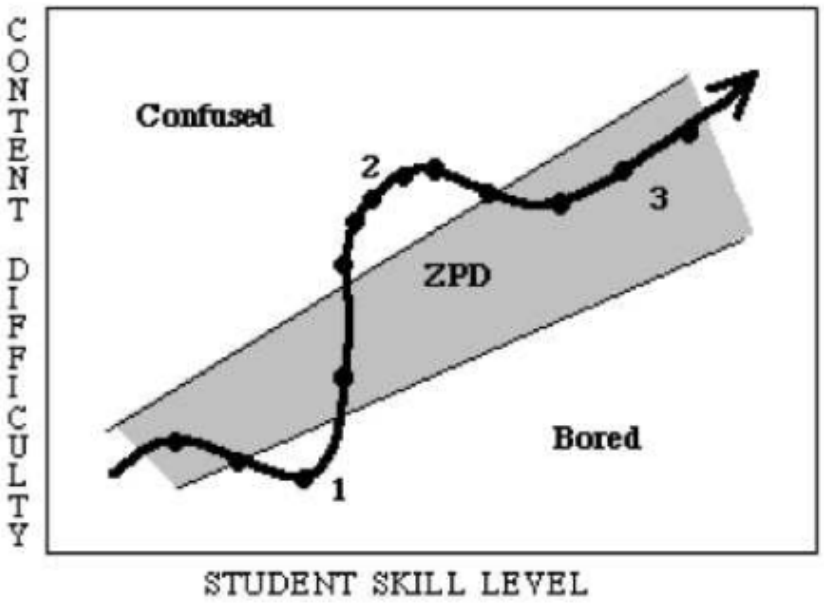
- Introduction to ITS
- Why are ITS effective?
  - Feedback and hints on steps (not just answers)
  - Adaptive task selection
- Why are ITS uncommon?
  - Model development is costly
  - ITS disrupt classroom social structure
- Can crowd-sourcing help solve ITS's problems?

Next

The Zone of Proximal Development (ZPD)



(a) Optimal Game Play Corridor

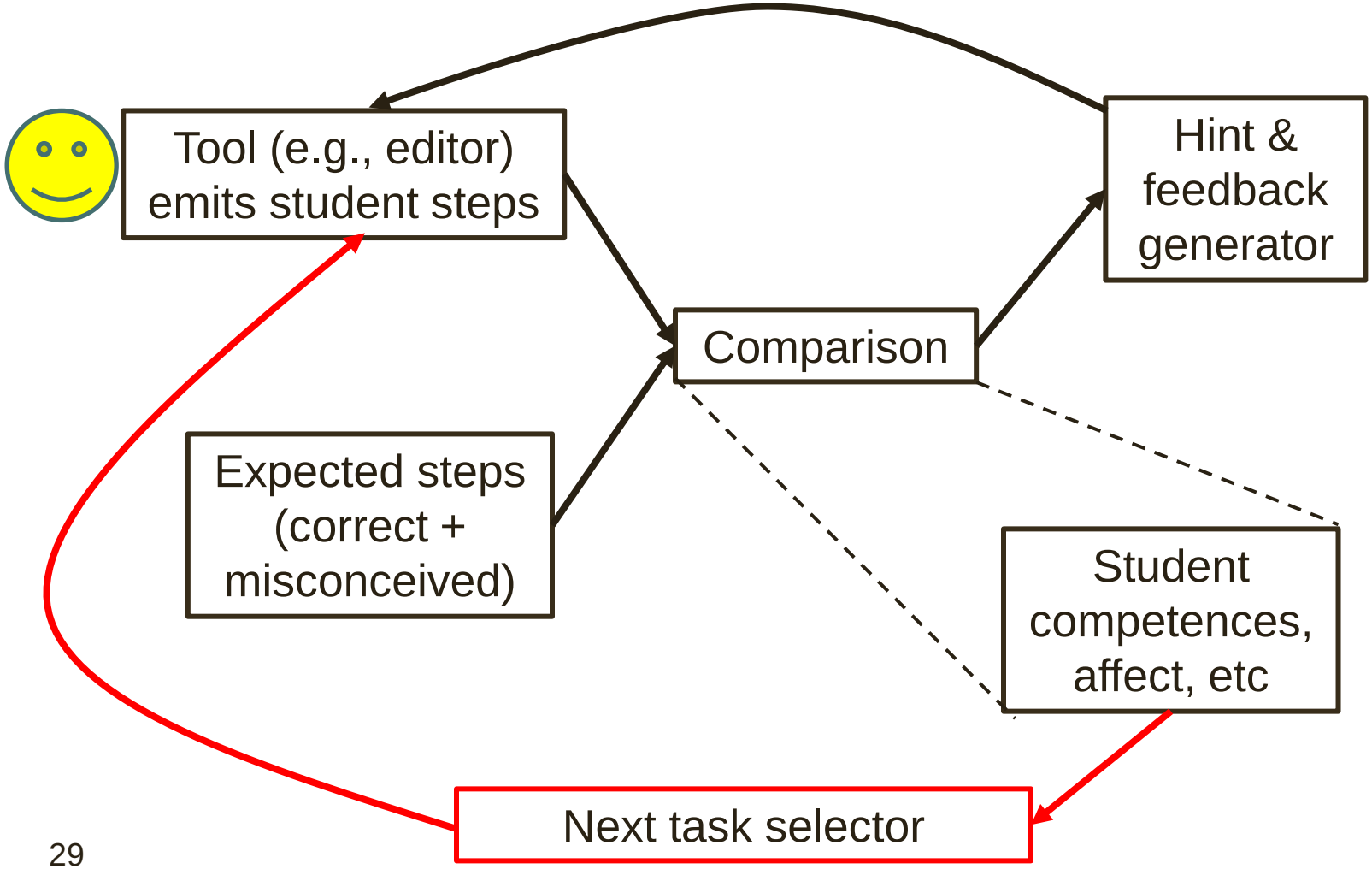


(b) Zone of Proximal Development

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

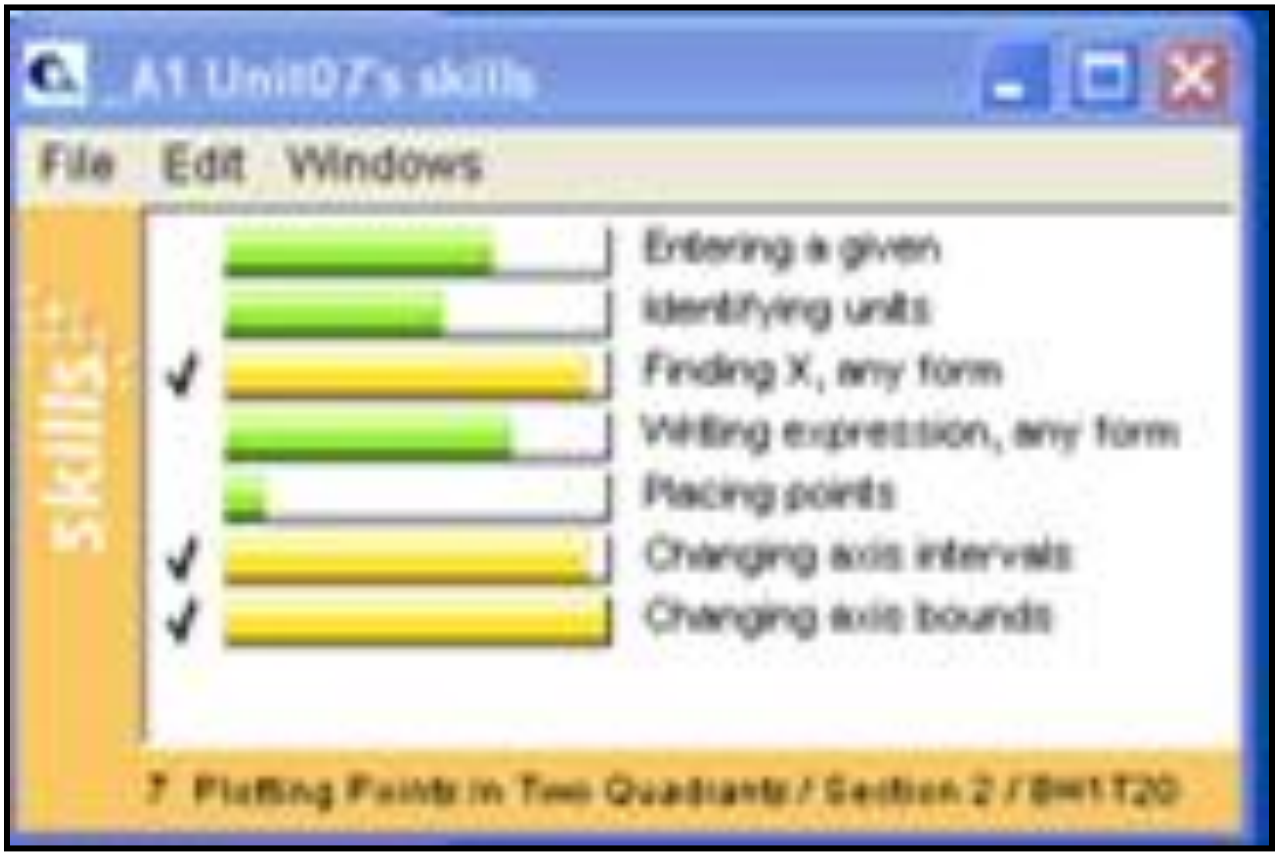
Mastery learning implemented by the red path below



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

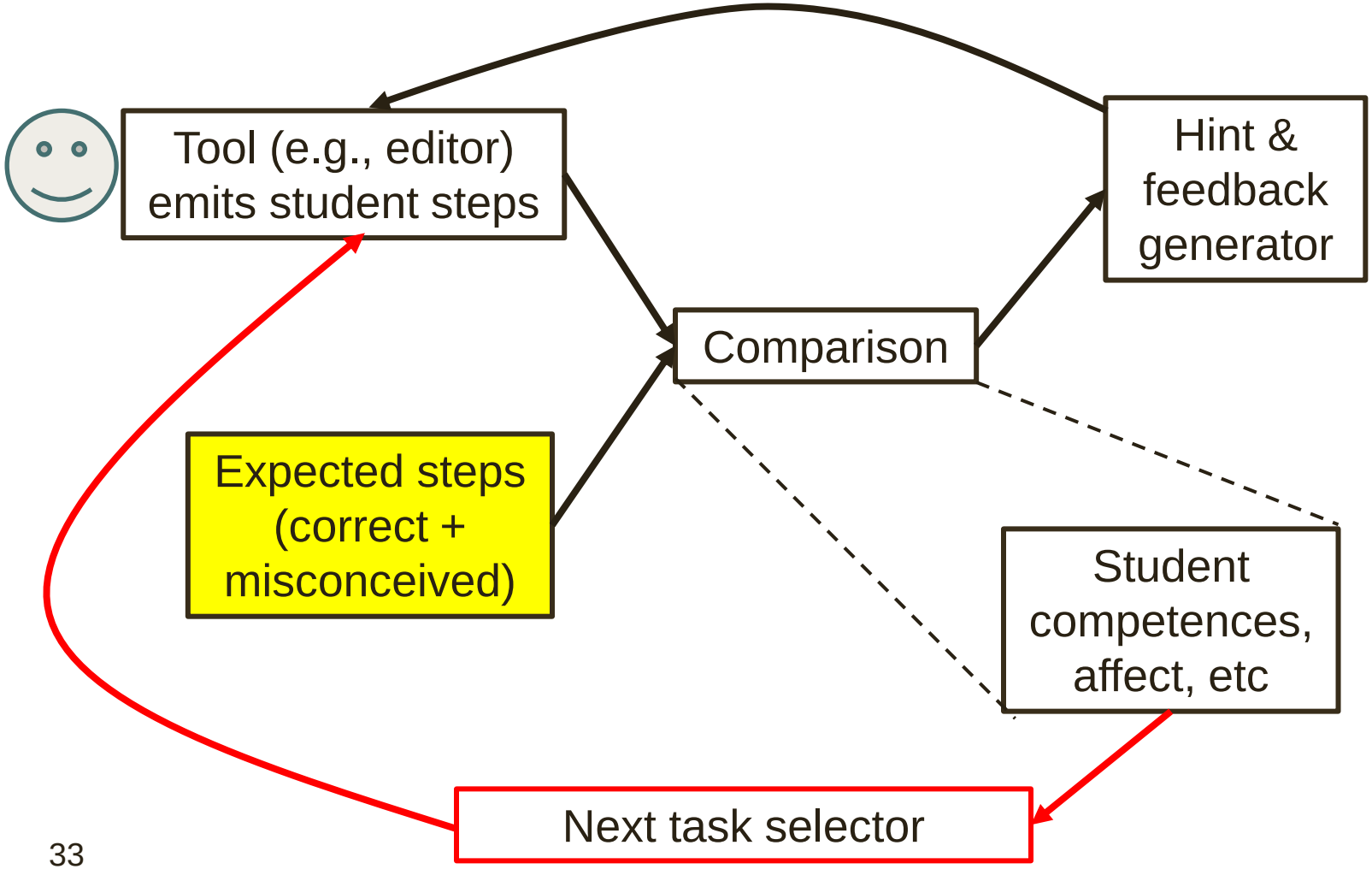
Assessment can be shown to student
(Carnegie Learning's Cognitive Algebra I Tutor)



Outline

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