Introduction to Tutorial Dialogue Systems

H. Chad Lane

CS599: Computational Approaches to Natural Language Dialogue Modelling

Plan

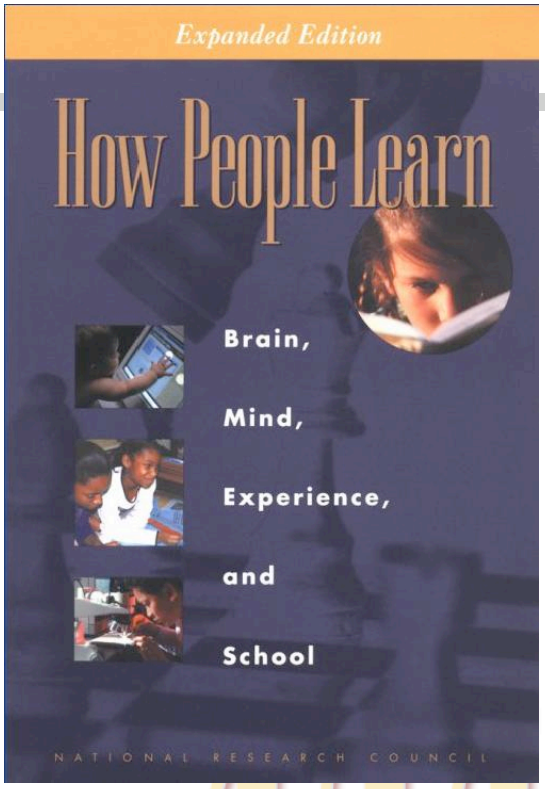
- tutoring
- existing systems
- results
- where it's happening & who's doing it
- open questions & future work
- web links & key reference points

Human learning

HPL highlights three key findings:

1. student preconceptions / incoming knowledge
2. to reach competence, a student must…
   a. have factual knowledge
   b. understand in context
   c. organize knowledge for retrieval & application
3. "metacognitive" approach to instruction
   - student control & self-monitoring

http://www.nap.edu/openbook/0309065577/html/



Bransford, Brown, Cocking (Eds.), National Academy Press, 2001.

Effective learning

Learning occurs when students…

- encounter obstacles and work around them
  - impasses / repair theory (VanLehn, 90)
- self-explain (Chi et. al. 89, 94, 01)
  - what worked / what didn't
  - how new information fits in to what is known
  - "mental" steps & reasoning behind actions

Human tutoring

- many motivations for tutoring
  - remediation of a deficiency
  - acquisition of new knowledge
  - learning of new skills
- "2 sigma problem" (Bloom, 1984)
  - the average student in 1-1 tutoring condition scored better than 98% of classroom students
- Why does tutoring work?
  - tutor-centered vs. student-centered vs. interaction

Intelligent tutoring

- application of AI to the problems of education
- (some possible) broad dimensions:
  - teaching (pedagogical) / performance (homework help)
  - product / process
  - dialogue / gui-based interaction
  - discovery / guided
- common approaches
  - rule-based / model-tracing (Anderson, 1995)
  - constraint-based modeling (Ohlsson 95)
  - plan recognition (Johnson, 1995)

ITS architecture

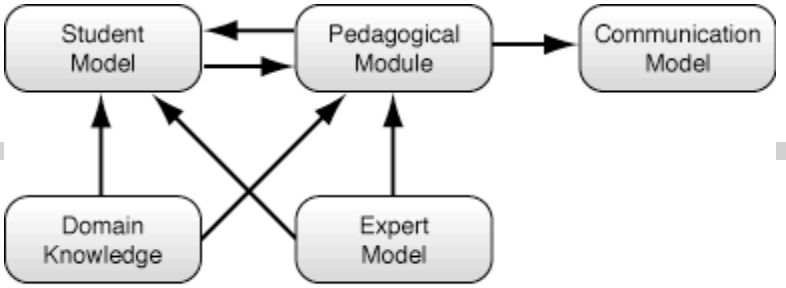


Figure 1: Interactions of components in an intelligent tutoring system.

Typically there are 4 modules:

- **interface / environment:**
  - simulated paper, simulated device, etc.
- **expert model:**
  - problem solver, ideal student, recognizer
- **student modeller:**
  - representation of student's knowledge state
  - plan recognition / track current knowledge
- **pedagogical module:**
  - compares user actions to expert model
  - decides when and how to intervene
  - generates feedback (hints, questions, pumps, etc.)

Closing the gap

Kind of tutor	effect size	Reference
"Expert" human tutors	2.0	Bloom 84
Best ITSs	1.0	Anderson 95 Anderson & Koedinger 97 VanLehn 01, 05
Computer-Aided Instruction	.42	Niemiec & Walberg 87
Inexperienced tutors	.4	Cohen, Kulik, & Kulik 82

What is different about human tutoring?

How can we "close the gap"?

Why natural language?

- NL offers indirect techniques for interaction:
  - signaling disagreement or uncertainty
  - giving positive feedback/approval
  - switching topics
  - taking/relinquishing initiative
- Opens door on more advanced tutoring techniques:
  - directed lines of reasoning (& multi-turn dialogue strategies)
  - elicitation/generation > recognition
  - scaffolded self-explanation
  - entrapment (the original meaning of "Socratic")
  - abstraction & concept-focused interaction

Graeser's 5-step frame

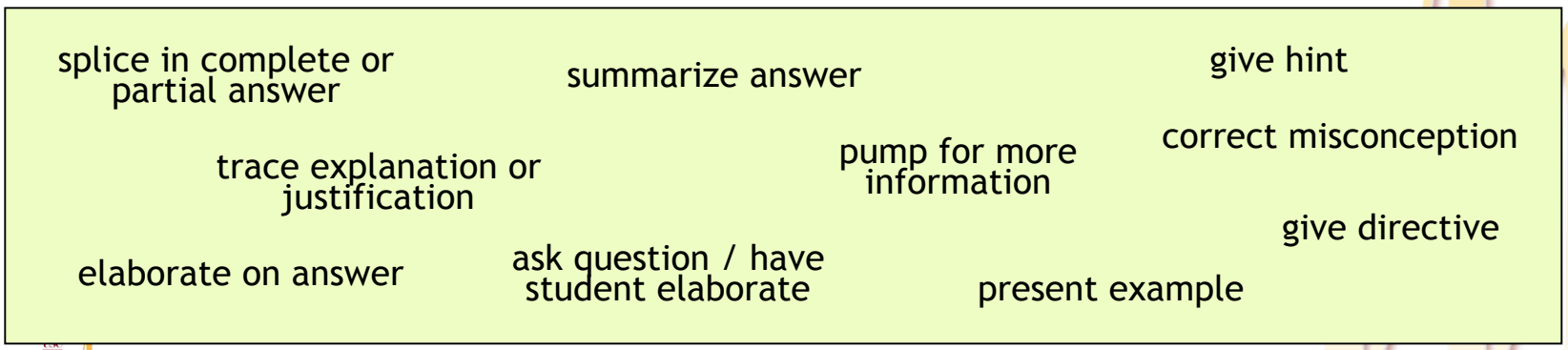
1. Tutor asks question (or presents problem)
2. Learner answers question (or begins to solve problem)
3. Tutor gives short immediate feedback on the quality of the answer (or solution)
4. The tutor and learner collaboratively improve the quality of the answer.
5. The tutor assesses the learner's understanding of the answer

Happens in classroom instruction

Happens in 1-1 tutoring

Step 4: improving the student's answer

- When asking why tutoring is better than classroom instruction, the focus has to be on steps 4 & 5.
  - many believe step 4 is where the magic is
- What are examples of tutoring tactics for improving a student's answer?



Dialogue-based tutoring systems

- How to use dialogue?
  - teach
  - prepare → solve problems → review
- some of the earliest ITSs were dialogue-based:
  - SCHOLAR (Carbonell, 70)
  - WHY (Collins, 77)
  - SOPHIE (Brown, 73)
- but were limited by NLP technology & no one did pedagogical evaluations

Plan

- tutoring
- existing systems
- results
- where it's happening & who's doing it
- open questions & future work
- web links & key reference points

AutoTutor (Memphis)

- Textual conversation with an **animated agent tutor**
- Originally for Computer Literacy
  - also for Newtonian Physics and Research Methods
- Goal: get (long) answers to **general, concrete questions** and elicit/correct student knowledge
  - e.g.: 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.

AutoTutor

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.

Talking head
- Gestures
- Synthesized speech

- Presentation of the **question/problem**

Log of previous turns

moves in its orbit around the sun . Does the earth pull equally on the sun? Explain why.
Student:

Tutor: Is there anything you can add to this?
Student:

Tutor: Kind of.
Tutor:
Tutor: How does Newton's law of motion apply to this situation?
Tutor:

Type your response here:

Submit

Settings...

Dialog history with
- tutor turns
- student turns

Student input (answers, comments, questions)

AutoTutor: Dialogue

- Selects dialogue moves from:
  - Positive/negative feedback
  - Prompts
  - Hints
  - …
- Students can ask Wh- and Yes/No-questions
- Based on a "Dialogue Advancer Network":
  - FSM that selects the next move according to student's last utterance
- Latent Semantic Analysis to match student answers with expectations

AutoTutor cont.

- form filling approach to dialogue management
  - value of the slots authored ahead of time in a *curriculum script*
  - answer "aspects"
- Mostly **system-initiative**
  - no memory (if the student asks a question, the system forgets what it is doing)
- Global strategy fixed (by system architecture)

AutoTutor tutoring rules

- **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 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]

Sample AutoTutor dialogue:

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

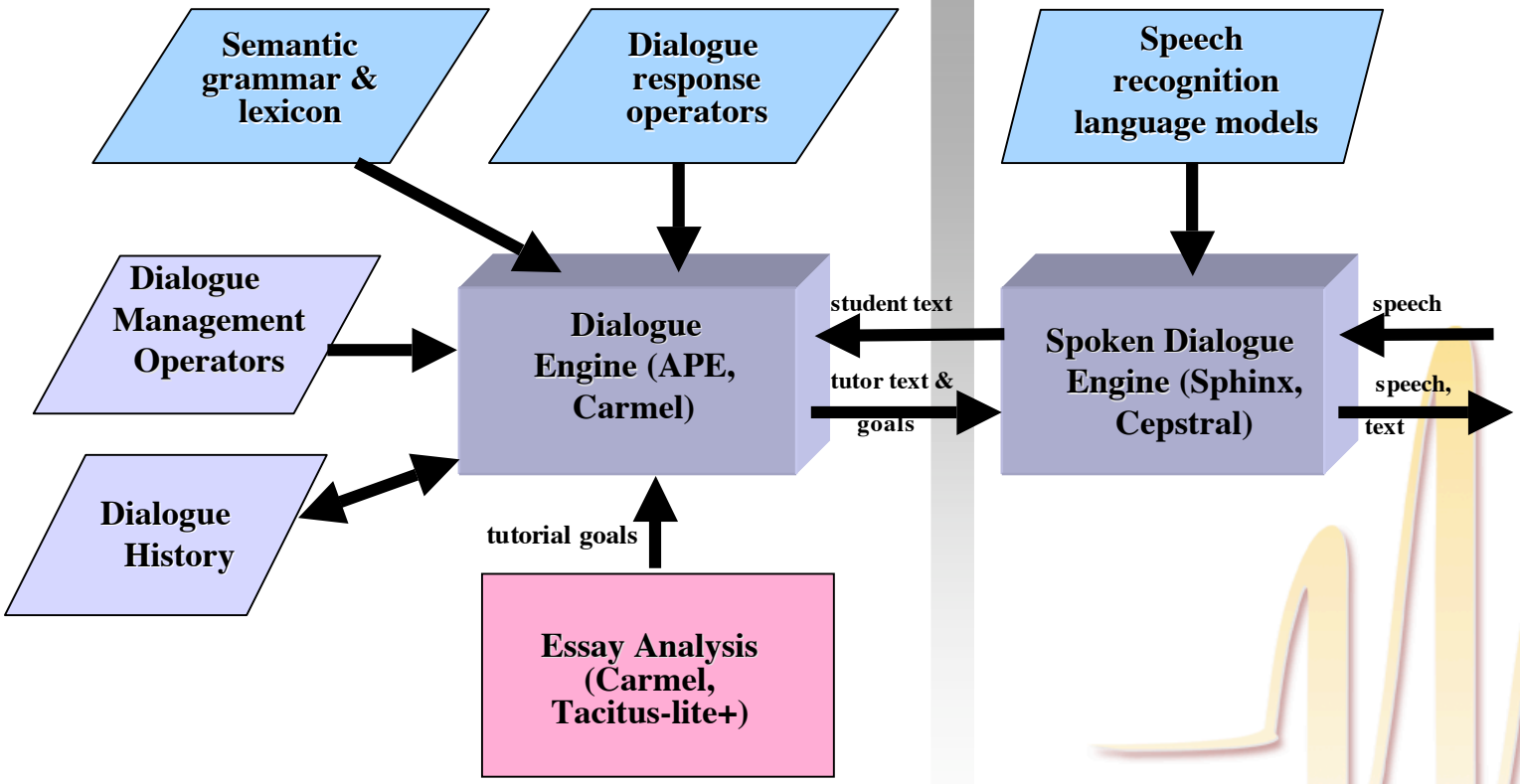
TUTOR-3: Yeah. For what type of connection do you need a network card?

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.

Dialogue Enabling Technologies (LRDC, Pittsburgh)



Why2-Atlas

ITSpoke

Atlas/Andes

- Andes: ITS for physics, no natural language
- Atlas: “add-on” to ANDES
  - provides **Knowledge Construction Dialogues** for hints (main task/evaluation is left to ANDES)
- KCD: recursive FSM
  - **Reactive planner** to pick next KCD
  - Can insert subdialogues (clarification, rectification…) and go back to original topic

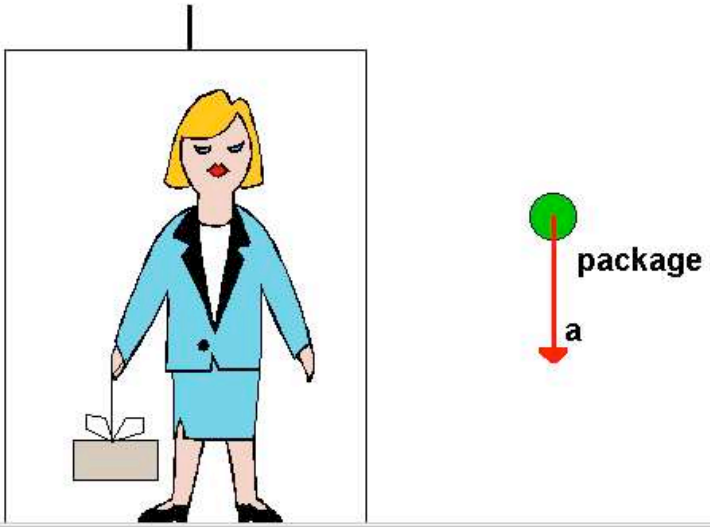
ANDES Physics Workbench - [Exdt1a-Solution.fbd]

File Edit Diagram Variable View Help

An elevator slows to a stop from an initial downward velocity of 10.0 m/s in 2.00 seconds. A passenger in the elevator is holding a 3.00 kilogram package by a vertical string.

What is the tension in the string during the process?

Answer:



Variables

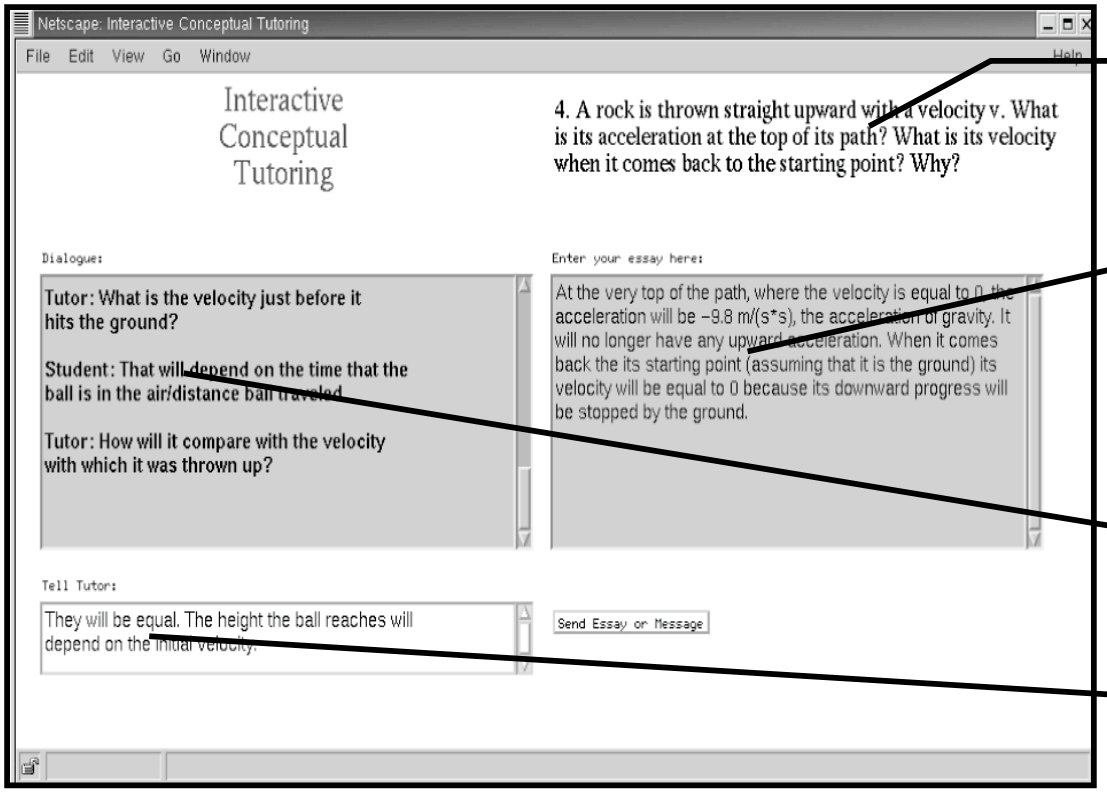
Name	Definition
T0	elevator at 10 m/s
T1	elevator at a stop
mpackage	mass of package
a	magnitude of the average Acceleration of package during T0 to T1

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.

T: What is the definition of acceleration?
S: I do not know
T: Let's look at it a different way. If I am driving north and you want to stop me which direction do you have to push?
S: north
T: But if the acceleration went the same direction as the velocity, then the elevator would be speeding up. Try again. If I am driving north and you want to stop me which direction do you have to push?
S: south
T: Right. What does that tell you about my acceleration?
S:

For Help, press F1 NUM 00:11:48

Why2/Atlas Qualitative Physics Tutoring System



A conceptual physics question

Student's essay: an answer to the question

Dialogue

Student types response here

How Atlas Works

- APE selects line of reasoning given Andes goal
  - An unsatisfied subgoal initiates tutor-student exchange
  - After exchange gets categorization of student response from CARMEL and identifies new subgoals
  - Initiates next subgoal (remedial or main line) until all subgoals satisfied
- CARMEL
  - Spell-checks and stems student input
  - Uses skipping feature of LCFlex to identify patterns and find associated answer category
  - Posts answer category for APE

A knowledge construction dialogue (KCD)

T: Let's consider the forces acting on objects after they have been thrown. As a runner pushes a ball away, what horizontal forces are acting on it while she is pushing it?

S: Gravity?

Subdialogue

T: In what direction does gravity act?

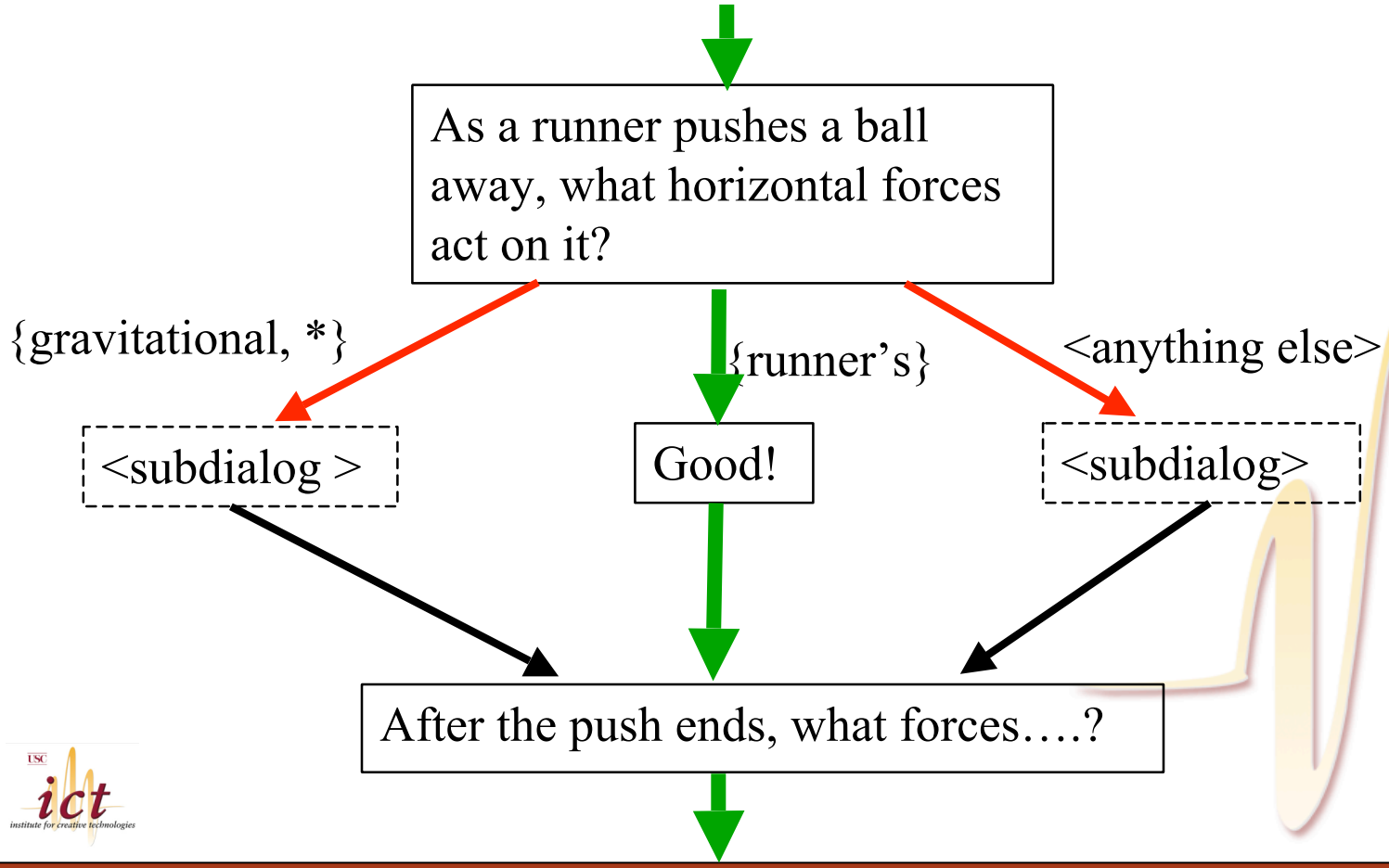
S: downward

T: Yes. So is a downward force a horizontal or vertical force?

S: It is vertical

T: So once again on the original question. As a runner pushes a ball away, what horizontal forces are acting on it while she is pushing it?

Push-down automata for low-level dialogue management



BEETLE (Edinburgh)

- domain: basic electricity and electronics
- attempting to integrate work from dialogue system community into tutoring
- Fully **plan-based tutorial dialogue**:
  - Top tier: global strategy/repair when failure
  - Middle tier: handles specific tasks according to the situation
  - Bottom tier: performs primitive dialogue actions
- not completely implemented yet

PACO: Pedagogical Agent for Collagen (USC/ISI, Mitsubishi, MITRE)

- **Simulation-based** training
- views tutoring as an instance of collaboration
- **Domain-independent**:
  - adapts to any simulator (e.g. Gas Turbine Engine)
- **Collaborative Discourse Theory-based**:
  - Rules describe interactions between three agents: student, tutor, simulator
  - Discourse acts: both utterances and domain actions
- plans used for both recognition and agent actions

example PACO dialogue

(1) Paco sets the state of the simulator.
(2) Paco "I've reset the simulation so that we can try engaging an engine under slightly different conditions."
(3) Paco "Let's engage engine one."
(4) Student "Ok."
(5) Paco "You take it from here."
(6) Student "Let's start the generator."
(7) Paco "Good."
(8) Paco (Pointing) "Now, notice that the alarm light is on."
(9) Paco "Before starting a generator, always make sure that the alarm light is off so that you'll know when a new alarm appears."
…

…
(37) Paco "You take it from here."
(38) Student opens fuel valve one on engine one.
(39) Paco "Good."
(40) Student opens the bleed valve on engine one.
(41) Paco "That was a correct action, but I wasn't expecting you to do it now."
(42) Paco "Let's go back to opening the fuel valves on engine one."
(43) Student opens fuel valve two on engine one.
(44) Paco "Great."
(45) Student presses the engage button on engine one.
(46) Paco "Whoops, you can't do that yet."
(47) Student "What next?"
(48) Paco (Pointing) "Press the on button on engine one."

Stanford's CSLI System

- Also **simulation-based** (Shipboard Damage Control)
- **Complex dialogue management:**
  - tree based activity model (similar to CMU Communicator) built dynamically
- **Separation between dialogue management and tutoring strategy:**
  - Tutoring Module constructs the activity tree using *recipes* while the Dialogue Manager uses the tree to conduct the dialogue

CIRCSIM (Illinois Inst. of Tech.)

- cardiovascular physiology: (Evens, 01)
  - how blood pressure is regulated in the human body
- student makes predictions based on changes
  - e.g., "Inotropic State and Heart Rate and Total Peripheral Resistance are directly affected by reflex. What is the correct value of TPR?"
- main components:
  - planner, text generator, NLU, student model
- a variety of issues addressed:
  - initiative, turn-taking, high-precision NLU, text planning, tutoring tactics, hedging, multi-model interaction, analogies, ontologies

Plan

- tutoring
- existing systems
- results
- where it's happening & who's doing it
- open questions & future work
- web links & key reference points

Does NLT work?

- typical comparison: NLT vs. reading
- AutoTutor:
  - improvement of .4 - .6 standard deviations
- Andes/Atlas:
  - mixed: some wins, some ties
  - win vs. hint-sequences (Rose, Jordan 01)
  - wins with reflective tutoring (Katz, 03)
  - ties in qualitative physics (VanLehn, in prep.)
  - (note) Andes alone was very successful (VanLehn, 05)

Does NLT work? (cont.)

- CIRCSIM:
  - tie on factual knowledge
  - wins on predictive abilities & number of misconceptions
- ProPL – novice programming tutor: (Lane, 05)
  - ties on planning, but wins on tacit knowledge
  - students more confident in read-only condition
    - "illusion of knowing"
- Geometry explanation tutor: (Aleven, 04)
  - NL > menu-selection & canned fbk

Is the gap closing?

- *in some cases, but not all*
  - in an unpublished meta-analysis by VanLehn:
    - 19 studies showed learning gains where NLT > Control on at least one achievement measure
    - 12 studies showed NLT = Control in 12 studies
    - *it seems NLT > reading alone*
    - *but NLT = reading + problem solving*
- need more ITS vs. ITS w/dialogue comparisons

Plan

- tutoring
- existing systems
- results
- where it's happening & who's doing it
- open questions & future work
- web links & key reference points

Who is doing NLT research?

- Pittsburgh:
  - Pitt: VanLehn, Litman, Katz, Jordan, Pappuswamy, …
  - CMU: Rose, Aleven, Koedinger, Levin, …
- Memphis: Graesser, Person
- Stanford: Peters, Clark
- Edinburgh: Moore, Zinn
- USC: Johnson, Traum, Core, Lane
- IIT: Evens, Michael, Rovick
- (many, many, more…)

Plan

- tutoring
- existing systems
- results
- where it's happening & who's doing it
- open questions & future work
- web links & key reference points

Some open questions

- What is the required level of NLU sophistication?
  - statistical vs. high-precision NLU
  - pros & cons of LSA
    - easy to train
    - but misses some important ideas (such as negation)
- What techniques from the dialogue community can/should be used by NLT system developers?
  - BEETLE
  - PACO also an example (but of ITS & CDS)

More open questions

- What does it mean to be an "expert" tutor?
  - and other pending issues from Ed. Psych. research
- Will improved NLP/dialogue systems lead to improved learning in tutorial dialogue systems?
  - sometimes deep understanding is not necessary
  - students seem OK with hiccups
- How can we encourage longer student utterances?
  - even in the face of NLU failures?
  - is this a good idea? What else should be encouraged?

Ongoing & future work

- authoring tools (Pitt, Edinburgh, CMU)
  - major focus in ITS and NLT communities
  - desire non-experts to create tutor utterances and expected answers
- multi-modal interaction
  - referring to diagrams & simulations in dialogue
  - e.g., PACO pointing (Steve did this as well)
- studying tutoring expertise (Person @ Memphis)
- simulated peers/students
- tutorial dialogue in immersive environments

ICT
institute for creative technologies

Plan

- tutoring
- existing systems
- results
- where it's happening & who's doing it
- open questions & future work
- web links & key reference points

sources

- Some slide content borrowed & modified from:
  - Kurt VanLehn (www.cs.pitt.edu/~vanlehn)
  - Neil Heffernan (http://web.cs.wpi.edu/~nth)
  - Johanna Moore (http://www.hcrc.ed.ac.uk/~jmoore/)
  - Pam Jordan (http://www.pitt.edu/~pjordan/)
  - Maxim Makatchev (http://www.pitt.edu/~maxim/)
  - CMU dialogues on dialogues group

Links to research groups:

- AutoTutor/Memphis:
  - http://www.autotutor.org/
- CIRCSIM/IIT:
  - http://www.cs.iit.edu/~circsim/
- Pitt:
  - http://www.pitt.edu/~circle/Projects/Atlas.html
  - Why2: http://www.pitt.edu/~vanlehn/why2000.html
  - Tutoring Scientific Explanations: http://andes3.lrdc.pitt.edu/ITR/tsed.html
  - TuTalk (Litman): http://andes3.lrdc.pitt.edu/TuTalk/

More links

- Edinburgh:
  - BEETLE: http://groups.inf.ed.ac.uk/beetle/
- USC
  - Tactical Language Tutor (ISI): http://www.isi.edu/isd/carte/proj_tactlang/
  - Reflective Tutoring & XAI (ICT):
    - http://people.ict.usc.edu/~lane/
    - http://people.ict.usc.edu/~core/