Background

• Developing diagnostic expertise (for example in Pathology) is difficult and time-consuming

• In domains outside of medicine, intelligent computer-based training is common
  – Aviation simulators
  – Nuclear and power plant simulators
  – Military

• Research in many domains has shown that computer systems can simulate the well-known benefits of one-on-one teaching
Outline

• What are intelligent tutoring systems (ITS)?
• How did we develop SlideTutor?
• How does SlideTutor work?
• How effective is the system?
• What are our plans for future research?
Q: What are Intelligent Tutoring Systems?
Intelligent Tutoring Systems

- Adaptive, flexible, *individually tailored* instruction
- Not ‘*text and test*’ but rather *coached practice environments*
- System able to ‘solve the problem’ on its own and therefore able to provide feedback on student actions
- Monitor student’s progress and change teaching based on how the student is learning
A Ballroom Dance Lesson

Student leads

Right step?  
...teacher follows

Wrong Step?  
...teacher corrects

Lost?  
...teacher leads
Intelligent Tutoring Systems

- Many, many successful systems in diverse domains
  - mathematics, programming, physics,
  - F15 fighter avionics troubleshooting,
  - proven educational benefit over classroom learning
- Very few Medical ITS have been developed – none have been evaluated:
  - GUIDON and GUIDON II (Clancey)
  - Cardiac-Tutor (Woolf et al) – ACLS
  - Rad-Tutor (Azevedo, Lajoie)–Mammogram Interpretation
Student Model

- Collect data on what student does
- Make predictions on what student knows
- Provide data for pedagogic decision making

Expert Module
- Allow correct steps
- Correct errors
- Give hints on next step

Pedagogic Knowledge
- Case sequence
- When to intervene
- How much to intervene
- How to intervene

Interface
Why are ITS so hard in Medical Domains?

• Expert Module
  – Uncertainty and missing information
  – Enormous amounts of declarative knowledge
  – Knowledge changes over time, KB requires verification/maintenance
  – Carefully selected and controlled case bases needed

• Interface
  – No formal problem-solving notation
  – Hard to reproduce environments in which medical PS occurs

• Student and Pedagogic Models
  – Unclear how to model acquisition of skills which involve so much declarative knowledge. How decomposable are these skills really?
  – Very little research to guide pedagogic modeling outside of causal reasoning domains (CircSim)

• Evaluation and Deployment
  – Limited access to subjects for formative and summative evaluation
  – No classrooms, potentially hard to deploy systems when they are done
Q: How did we develop SlideTutor?
Data Collection
Methods

- Studied Novices, Intermediates, and Experts
- Think aloud protocols
- People can articulate the intermediate steps of cognition that are available in Working Memory
- Collect detailed data and analyze
- Develop ideas about how expertise is acquired
Research Paper

Development of Visual Diagnostic Expertise in Pathology: An Information-processing Study

Rebecca S. Crowley, MD, MS, Gregory J. Naus, MD, Jimmie Stewart III, MD, Charles P. Friedman, PhD

Abstract Objective: To identify key features contributing to trainees' development of expertise in microscopic pathology diagnosis, a complex visual task, and to provide new insights to help create computer-based training systems in pathology.

Design: Standard methods of information-processing and cognitive science were used to study diagnostic processes (search, perception, reasoning) of 28 novices, intermediates, and experts. Participants examined cases in breast pathology; each case had a previously established gold standard diagnosis. Videotapes correlated the actual visual data examined by participants with their verbal "think-aloud" protocols.

Measurements: Investigators measured accuracy, difficulty, certainty, protocol process frequencies, error frequencies, and times to key diagnostic events for each case and subject. Analyses of variance, chi-square tests and post-hoc comparisons were performed with subject as the unit of analysis.

Results: Level of expertise corresponded with differences in search, perception, and reasoning components of the tasks. Several discrete steps occur on the path to competence, including development of adequate search strategies, rapid and accurate recognition of anatomic location, acquisition of visual data interpretation skills, and transitory reliance on explicit feature identification.

Conclusion: Results provide the basis for an empirical cognitive model of competence for the complex tasks of microscopic pathology diagnosis. Results will inform the development of computer-based pedagogy tools in this domain.

### Acquisition by Skill Type

<table>
<thead>
<tr>
<th>Search and Detection</th>
<th>Feature Identification</th>
<th>Feature Refinement</th>
<th>Hypothesis Triggering</th>
<th>Hypothesis Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor search skills</td>
<td>Very limited ability to recognize and name histopathologic features</td>
<td>Features almost never refined</td>
<td>Few hypotheses triggered</td>
<td>Very limited ability to connect features to hypotheses</td>
</tr>
<tr>
<td>Limited perceptual recognition of areas of interest</td>
<td>27,34,35,36</td>
<td>56</td>
<td>Triggered hypotheses are overly general - lack knowledge of diagnostic space</td>
<td>Lack knowledge base needed for backwards reasoning</td>
</tr>
<tr>
<td>Accurate search skills</td>
<td>Many features identified, poor filtering</td>
<td>Many hypotheses triggered, very early in diagnostic process</td>
<td>Error prone in connecting features to hypotheses</td>
<td></td>
</tr>
<tr>
<td>Accurate perceptual recognition of areas of interest</td>
<td>Frequent errors in feature identification</td>
<td>Hypothesis set overly broad</td>
<td>Backwards reasoning made necessary but also more challenging by broad hypothesis set</td>
<td></td>
</tr>
<tr>
<td>Accurate search skills</td>
<td>Express uncertainty even when feature identification is correct</td>
<td>Fail to change hypotheses even when confronted with inconsistent data</td>
<td>Accurate in connecting features to hypotheses</td>
<td></td>
</tr>
<tr>
<td>Rapid perceptual recognition of areas of interest</td>
<td>Few, highly salient features identified</td>
<td>Many hypotheses triggered when necessary</td>
<td>Backwards reasoning to differentiate among more focused set of hypotheses</td>
<td></td>
</tr>
</tbody>
</table>

### ITS Design Requirements

- **EARLY (1.1)** monitor location on slide and direct more novice students to areas of interest using symbolic and visual cues. (1.2) provide explicit strategies to compensate for poor search skills.
- **LATE (1.3)** monitor location on slide to ensure that ROI has been found but otherwise limit feedback on search and detection.

- **EARLY (1.4)** require that students connect visual features on the slide with feature names and provide feedback. (1.5) encourage complete articulation of features. (1.6) allow full exploration of feature space, but provide feedback on relevancy.
- **LATE (1.7)** allow students to “jump” to the diagnosis as long as the ROI has been seen. (1.8) encourage students to search for features in order most efficient for narrowing search.

- **EARLY (1.9)** encourage complete feature refinement and provide feedback.
- **LATE (1.10)** permit students to reason without fully refined features. (1.11) permit sequences in which new hypotheses require re-examination of feature refinement.

- **EARLY (1.12)** permit any hypothesis that is consistent with at least one piece of evidence. (1.13) encourage refinement and specialization of hypotheses.
- **LATE (1.14)** encourage hypotheses that are consistent with all of the features. (1.15) help students learn sets of hypotheses that share similar features. (1.16) permit sequences in which new features or feature refinements require re-examination of triggered hypotheses.

- **EARLY (1.17)** provide feedback on student generated relationships between features and hypotheses. (1.18) support both forwards and backwards reasoning but guide students to use backwards reasoning to help build mental model of diagnostic space.
- **LATE (1.19)** help students learn when hypotheses should be excluded as students further refine features. (1.20) support both forwards and backwards reasoning but guide more advanced students to use backwards reasoning more selectively.

**Table 1:** Developmental Cognitive Model of Visual Diagnostic Expertise and resulting ITS Design Requirements

Crowley and Medvedeva, AIM 2006
Scaffolding skill acquisition

• Searching skills
  – Prevent students from interpreting non-diagnostic areas
  – Require that students see the entire slide
  – Monitor and provide feedback on specific search skills
• Feature identification
  – Encourage identification and full specification of evidence
  – Correct errors and give explanations
  – Combine visual and symbolic aspects of training
• Hypothesis triggering and testing
  – Provide feedback for forwards and backwards reasoning
  – Support efforts to distinguish between hypotheses
• But allow enough flexibility so that system will support students who are in many stages of skill acquisition
Q. How does SlideTutor work?
Hints are hierarchically structured, providing increasing help.

System provides a hint on the next best step.

1. Student indicates a new finding.
2. Clicks on the finding, creating an 'x'.
3. Selects the finding.

Tutor permits the correct answer.

Student requests a hint about the next best step.

Later, the hint tells the student the important quality and then opens a menu to display choices.

1. Student indicates a new hypothesis.
2. Selects the hypothesis.

1. Student indicates a new refute link.
2. Student draws between the evidence and the hypothesis.

Object flashes.

Bug message explains the error.

Actually, blister supports acute burn.
Cluster does not support Acute Burn

A hypothesis… reasonable but wrong

"Backwards" reasoning
Analysis by hint and error

![Graphs showing hint and error rates for Case-Focused and Knowledge-Focused approaches across different problems.](image)
Analysis by learning curves

![Graphs showing normalized count over trials for different conditions.](attachment:image.png)
Q. How effective is the system?
Subjects

- 21 pathology residents from 2 academic programs
- PGY1 – PGY5
Learning gains on multiple-choice test

Multiple Choice Test Scores

Score

Pretest | Posttest | Retention

Test

0.0

20.0

40.0

60.0

80.0

100.0

***

***
Learning gains on case diagnosis test

Case Diagnosis: Tutored Patterns vs Untutored Patterns

![Bar chart showing learning gains on case diagnosis test](chart.png)

- **Pretest**
- **Posttest**
- **Retention**

Scores are indicated on the y-axis, with tests labeled on the x-axis. The chart compares tutored patterns and untutored patterns, with statistical significance marked by ***.
Meta-cognitive gains – What do they know about what they know?

Crowley et al, AIED 2005
Tutoring other cognitive skills

Saadawi et al, AHSE (in press)
Learning Gains

Saadawi et al, AHSE (in press)
Conclusions

• With some critical modifications, the ITS paradigm can be used effectively in medicine

• SlideTutor significantly improves performance in diagnostic problem solving and reporting, even after a single session, and learning gains are retained at one week (diagnostic)
Q. What are our plans for future research?
Ongoing and future Work

• Deployment
  – Develop case library and additional knowledge models
  – Cancer Training Web – collaboration with University of Pennsylvania

• Evaluation
  – Summative evaluation of SlideTutor against “standard practice”
  – Can SlideTutor decrease medical errors? (Collaboration with D. Grzybicki, AHRQ)

• Basic Research
  – Student Modeling and performance prediction
  – Meta-cognitive Training
  – Comparing pedagogic approaches (depth v. breadth)
Acknowledgements

- **SlideTutor Project**
  - Collaborators:
    - Drazen Jukic, MD, PhD
    - Gilan Saadawi, MD, PhD
    - George Xu, MD, PhD (University of Pennsylvania)
    - Dana Grzybicki, MD, PhD
  - Programmers:
    - Olga Medvedeva
    - Eugene Tseytlin
    - Girish Chavan
  - Research Associates:
    - Elizabeth Legowski
    - Melissa Castine
  - Students:
    - Michael Yudelson
    - Velma Payne
  - Funding:
    - *National Library of Medicine 1R01 LM007891*
    - *National Cancer Institute 1R25 CA101959-and 2R25 CA101959*
    - *Agency for Health Care Research and Quality I UL1 RR024153-01(Grzybicki, PI)*
    - *National Library of Medicine Training Grant 5-T15-LM07059*
    - *University of Pittsburgh Competitive Medical Research Fund*