

# *Building from Cognitive Data: A Model of Visual Diagnostic Expertise in Pathology in Service of an Intelligent Tutoring System*

*Rebecca Crowley, MD*

*Center for Biomedical Informatics*

*University of Pittsburgh*



# Overview

- Background
- Cognitive analysis of expertise in Pathology
- Design/development of an Intelligent Tutor in Pathology



# Background

- Great expectations of Computer-Assisted Instruction (CAI) – not just in Medicine – have not been met.
- Cognitive science, AI, and education researchers developing Intelligent Tutoring Systems in other domains.
- Evaluation results of these systems are promising.
- Few systems of this kind have been developed in medical domains.



# *Why An Intelligent Tutor in Diagnostic Pathology?*

- A difficult task to learn, residents start as complete novices.
- Many rare patterns must be learned.
- Limited exposure to parts of the domain.
- Necessary technology becoming available – digital archives, whole-slide-imaging, “virtual microscopes”.



# *Intelligent tutors are different*

- Opportunity to build factual knowledge AND learn how to APPLY that knowledge.
- Make implicit aspects explicit and visible.
- Highly interactive - give guidance as student works by identifying and prevent errors in the intermediate steps.
- Flexible and adaptive curriculum – sequence of problems dependent on student’s development.
- A virtual apprenticeship - “scaffolding”.



# *Cognitive data is necessary to build an intelligent tutor*

- Every domain/task is different.
- Expert model
- Student model
- Interface





*A Cognitive Analysis of  
Expertise in Diagnostic  
Pathology*



# *Research Objectives*

- Identify the cognitive processes that expert pathologist's use.
- Describe how novices and intermediates are different.
- Characterize the kinds of errors made by novice and intermediate subjects, and their frequencies.
- Use this information to develop a teaching system in this domain



# *Information-Processing*

- Newell and Simon methodology
- Gain access to the cognitive processes by having subjects verbalize their thoughts as they perform a task (verbal protocols).
- Trace processes, aggregate and compare across level of expertise.
- Use protocols to identify errors.
- Verbalization doesn't alter performance, if you give the right kinds of instructions.



# *Cases & Subjects*

- Standardized set of cases in breast pathology
- 10 Novices – 3<sup>rd</sup> year medical students
- 10 Intermediates – 2<sup>nd</sup> and 3<sup>rd</sup> year residents
- 10 Experts – Attending pathologists with >10 years experience



# Data Collection



# *The data*

- ~29 hours of digital video
- 117 coded verbal protocols containing 7079 coded verbal segments (~90% complete)



# Coding the protocols

1 **Subject 1 – VS/Case 1**  
2  
3 Total time: 8:09  
4 Time until final diagnosis suggested: 3:36  
5  
6 Holds slide up to light (0:22):  
7 This is a great big biopsy of something.  
8 It looks like it might be a prostate at this power.  
9 And even looking... just holding the slide up to the light.  
10 I'm seeing some rather scary looking blue areas.  
11 Right out near the capsule.  
12 I'm getting a very early warning about this slide.  
13 Let's see what it looks like under the scope.  
14  
15 4X (2:13):  
16 Starting out at low power.  
17 Let's see. So we scan around a little bit.  
18 It turns out that my... Huh.  
19 It is the prostate.  
20 The architecture was so distorted  
21 that it took me a while of scanning around to be certain  
22 that I was indeed looking at prostate.  
23 What finally gave it away to me was that was within some of the ducts I'm seeing  
24 laminated concretions that are typical of prostatic tissue.

Transcribe verbatim

**Data Examination:**  
Identify-anatomic-location  
Identify-histologic-cue  
Identify-historical-cue

**Data-exploration and explanation:**  
Integrate-findings  
Note-absent-findings

**Data Interpretation:**  
Statement-of-hypothesis  
Confirm-with-present-finding  
Disconfirm-with-absent-finding  
Reject-hypothesis  
Search-discriminator

**Control Processes:**  
Assess-barriers-to-progress  
Evaluate-declarative-knowledge  
Recall-diagnostic-categories

Blank slide
2 Blank slide
3 Blank slide
4 Blank slide
5 Blank slide
6 Power change 70 Power change 25
7 Power selection low power
8 Not noted
9 What Value Identify anatomic location fibrous tissue
10 What value Identify histologic feature
11 Time 0:16
12 Goal Sub-goal identify specimen on this slide
13 What To Value Identify pathologic finding cellular infiltrate diffuse aspect of liver
14 When To Relationship Compare other areas patient with liver cellular
15 What To Relationship Compare other areas current area with fibrous tissue
16 What Value Quality priority Identify pathologic finding vessels little common
17 What Value Describe specific cellular organization
18 Assign significance Estimated versus hyperplastic changes
19 What To Relationship Compare current area other areas liver cellular
20

Develop coding scheme

Code protocols



# *An example protocol: An expert diagnosing DCIS*

## Expert E7:

- 3 Okay let's look at low power
- 4 I think the tissue is breast → *identify-anatomic-location*
- 5 I recognize some normal
- 6 Here is in situ carcinoma → *statement-of-hypothesis*
- 7 I have to find out if there is any invasion → *set-goal-identify*

# *An example protocol: An intermediate diagnosing DCIS*

## Intermediate I5:

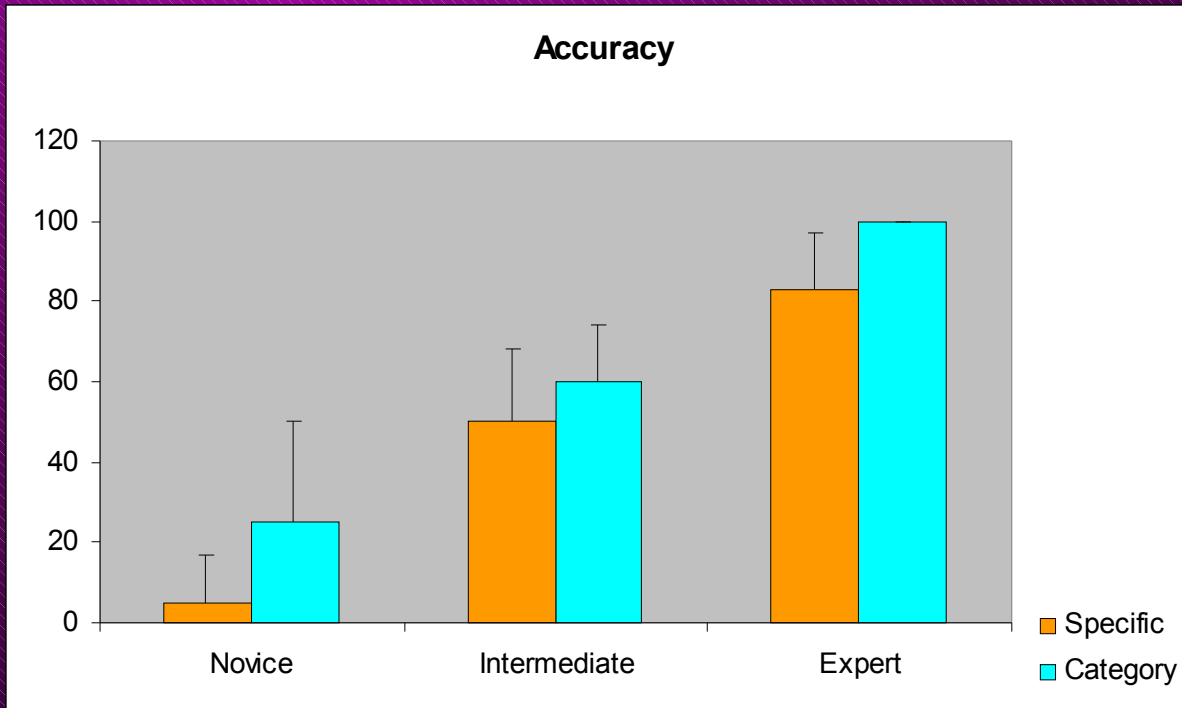
- 20 and some of the ducts that are expanded with small cells → *identify-histologic-cue*  
4 with focal, possibly central, area of necrosis. → *identify-histologic-cue*  
5 So just scan this slide around and try to determine some  
focal areas that I want to concentrate and focus on.  
6 Now I'm looking at some of the ducts that are expanded.  
25 And some of these ducts, they also have holes, → *identify-histologic-cue*  
26 and these are sort of punched-out holes, → *identify-histologic-cue*  
27 very uniform, which... → *identify-histologic-cue*  
29 So at this magnification, I think it is a DCIS. → *statement-of-hypothesis*

# *Processes & Themes*

- Search, Identification, Interpretation
- Transition to accurate, and then automatic searching
- Evolution of visual efficiency
- Shift from explicit feature identification and inference to rapid implicit pattern matching
- Certainty about evidence lags accuracy about evidence
- Development of goal-structured search and discrimination



# Overall accuracy



# *Identifying cognitive errors*

**With combined video and protocol analysis, it is often possible to trace an incorrect diagnosis to a very specific cognitive error:**

- **A case of LCIS misdiagnosed because the anatomic structure itself was misclassified.**
- **A case of carcinoma where the correct diagnosis was made on an area of normal tissue.**





*Design and development of  
an Intelligent Tutoring  
System in Pathology*



# *Building from cognitive data*

This is a multi-step process requiring three different sets of skills.

Searching the slide is a difficult first step.

Acquisition of skill often accompanied by uncertainty in classifying evidence.

Rapid pattern matching is preceded by explicit inference.

Model these processes using a model-tracing approach.

Include rules for searching and magnification use; connect the model to a virtual microscope.

Require the student to explicitly relate image features to diagnostic criteria. Give feedback.

Make these explicit reasoning steps visible to the student, require that they argue for and support their Dx in a reasoning interface.



# Model-tracing tutors

- One type of Intelligent Tutor
- Production-rule model of expert built in.
- System dynamically creates the student model.
- Every action in the interface mirrors an intermediate cognitive step.
- *Prevent mistakes* with immediate feedback – because each user action must match a fire-able sequence of rules or it is not allowed.
- Many different paths can be allowed.
- *Offers hints* by stepping ahead in the model.
- *Spot reasons for mistakes* and point them out to student – “Buggy rules”.
- Employ sequences that builds skills, and target the skills that have not been mastered for additional instruction.



# *A prototype model-tracing tutor in Pathology*

- Production-rule expert model developed from the protocol analysis.
- Tutor Development Kit (TDK) – Lisp environment for ITS development.
- A test domain – Proliferative epithelial lesions of breast.
- Interface building on existing technologies for whole-slide capture and web-based display.
- Model makes inference from symbolic information that must be coded for each slide (no machine vision).



# *Future Work*

- Complete protocol analysis – final data aggregation, statistical testing .
- Instantiate the tutor in a specific sub-domain of pathology – add to generic rule-set and build knowledge representation (WMEs).
- Connect the model to existing “virtual microscope” system .
- Develop knowledge-tracing and sequencing



# *Acknowledgements*

- Charles Friedman, PhD
- Michael Becich, MD, PhD
- Gregory Naus, MD
- Ken Koedinger, PhD
- Michael Sendek, Yukako Yagi, Susan Leeds
- National Library of Medicine



# Rule-Set

## Observation

Begin-Search  
Focus-Search-Next-Location  
Search  
New-magnification  
Check-Done

## Identification

Note-Informative-Area  
Identify-Present-Findings  
Set-Goal-ID-Absent-Findings  
Identify-Absent-Findings

## Interpretation

Generate-Hypothesis  
Support-Hyp-With-Present-Findings  
Support-Hyp-With-Absent-Findings  
Refute-Hyp-With-Present-Findings  
Refute-Hyp-With-Absent-Findings  
Ready-Reject-Hypothesis  
Reject-Hyp-Mutually-Exclusive  
Reject-Hyp-Insufficient-Evidence  
Accept-Hypothesis  
Make-Diagnosis-One-Entity  
Make-Diagnosis-Same-Entity  
Make-Diagnosis-Not-Mutually-Exclusive  
Make-Diagnosis-Mutually-Exclusive



# Interface design

The screenshot displays the PathTutor software interface, which is designed for medical education and diagnostic reasoning. The interface is organized into several key sections:

- Virtual Microscope:** Located on the left, it features a thumbnail labeled "SP-124" and a large, detailed histological image of a breast tissue section.
- Clinical History:** A text box on the right provides the patient's background: "Excisional biopsy of left breast from 56 year old female with areas of abnormal calcification on mammogram, not seen on previous exam."
- Hints:** A section below the clinical history, currently empty.
- Diagnostic Reasoning Diagram:** A central workspace showing a flowchart. It starts with two boxes labeled "DCIS" and "EH". A green arrow points from "DCIS" to an oval labeled "Ducts expanded". Another green arrow points from "EH" to the same "Ducts expanded" oval. Below this, a second oval labeled "Comedo necrosis" is shown, with a green arrow pointing from the "Ducts expanded" oval to it.
- Case Tree:** On the right side, a hierarchical tree structure shows the diagnostic process. It includes sections for "Location 1" (DIAGNOSIS - DCIS), "Location 2" (DIAGNOSIS - EH), and "Location 3" (HYPOTHESIS - DCIS and HYPOTHESIS - EH). The "Ducts expanded" hypothesis under "HYPOTHESIS - EH" is highlighted in blue.
- Navigation and Action Buttons:** At the bottom, there are eight buttons: "Show History", "Support Link", "Accept Hypothesis", "Local Diagnosis", "Show Hint", "Refute Link", "Reject Hypothesis", and "Global Diagnosis".



# *Classifying errors across the spectrum of expertise*

Search	Never get lesion into view
	Repeated search of uninformative areas
Identification	Lesion traversed but not noticed
	Finding misidentified
	Normal structure misidentified
Interpretation	Assign wrong significance
	Use wrong discriminator
	Insufficient evidence to accept hypothesis

