Artificial Intelligence Methods for Modeling and Assessing Collaborative Distance Learning

Amy Soller
Institute for Defense Analyses
asoller@ida.org

Mixture Models in Latent Variable Research
University of Maryland, Center for Integrated Latent Variable Research
May 19, 2006

The opinions or assertions contained herein are those of the author alone, and are not to be construed as official or reflecting the views of the Department of Defense, other U.S. Government agencies or departments, or the Institute for Defense Analyses.

Acknowledgments

- **Alan Lesgold, Ph.D.**
  - Professor & Dean, School of Education, University of Pittsburgh

- **Ron Stevens, Ph.D.**
  - Professor, Microbiology and Education, UCLA
  - Director of the IMMEX Project
Outline

“Artificial Intelligence (AI) Methods” for Modeling and Assessing
“Collaborative Distance Learning”

- Collaborative Distance Learning
- AI Methods: Intro to Hidden Markov Models (HMMs)
- HMMs for analyzing distributed online collaborative learning
- HMMs and Neural Networks for assessing student adoption of problem solving strategies
- Future Project Ideas

Collaborative Distance Learning

- Computer-supported collaborative learning tools enable learning across time & distance boundaries
  - On-line lectures
  - Lecture notes, & reference material
  - Tutorials & quizzes
  - Shared workspaces
  - On-line discussions - chat & bulletin boards
  - Group project management

Why?
The Collaborative Learning “Effect”

- Cognitive conflict (Doise, Mugny, & Perret-Clermont, 1975)
  - Conflicting viewpoints provoke justification, elaboration
  - Conceptual change (Roschelle & Teasley, 1993, 1995)
- Knowledge construction (Resnick, Levine, & Teasley, 1991)
  - Idea generation, explanation
  - Knowledge sharing → co-construction → learning
- Social skill development (Brown & Palincsar, 1989)
  - Mutual support
  - Cooperating, leading

When Teams Do Not “Function The Way They Ought To”

- Collaboration does not always enhance individual learning
  - Group size and composition – e.g. status, gender
  - Task structure
  - Rewards or incentives
  - Motivation
  - Communication and Cohesion
- CSCL technology falls behind in providing the kind of support that classrooms offer
Why Hidden Markov Models?

<table>
<thead>
<tr>
<th>Characteristics of Collaborative Learning Data</th>
<th>Characteristics of Hidden Markov Models (HMMs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group cognition develops (sequentially) over time</td>
<td>State-based model</td>
</tr>
<tr>
<td>Learning conversations take on a noisy, unpredictable character Peers interrupt each other, have multiple agendas of conversation</td>
<td>Stochastic method</td>
</tr>
<tr>
<td>Students achieve proficiency at different rates</td>
<td>Appropriate for sequences of indefinite length</td>
</tr>
<tr>
<td>Training Data is expensive to obtain</td>
<td>Able to model patterns well even with a limited amount of training data</td>
</tr>
</tbody>
</table>

A Brief Introduction to HMMs

- Markov Chain → probabilistic finite state machine
  - Arcs describe probability of moving between states
  - \( P(\text{sequence}) = \prod \text{probabilities along path} \)

- Hidden Markov Model (HMM)
  - Process (states) cannot be directly observed
  - Observations (stochastically linked to states) describe process
  - Several different paths may produce the same output
Definition of an HMM:

\[ \lambda = (A, B, \Pi) \]

\[ \Pi = (\pi_i) : \text{initial state distribution} \]

\[ A = (a_{ij}) = P(q_t = S_j \mid q_{t-1} = S_i) : \text{state transition matrix} \]

\[ B = (b_i) = P(v_k \mid S_i) : \text{observation symbol probability distribution} \]
**Definition of an HMM:**

\[ \lambda = (A, B, \Pi) \]

\( \Pi = (\pi_i) \): initial state distribution

\( A = (a_{ij}) = P(q_t = S_j | q_{t-1} = S_i) \): state transition matrix

\( B = (b_j) = P(v_k | S_j) \): observation symbol probability distribution

**HMM Applications & Examples (1/2)**

*Example 1: Online Collaborative Learning*

States represent which students are likely to take initiative

Observations represent likely student communication or action

* Notional Example

\( A = (a_{ij}) = P(q_t = S_j | q_{t-1} = S_i) \): state transition matrix

\( B = (b_j) = P(v_k | S_j) \): observation symbol probability distribution
HMM Applications & Examples (2/2)

*Example 2: Online Chemistry Problem Solving

States represent problem solving strategies student is likely to apply

Observations represent likely student problem solving actions

\[ A = (a_{ij}) = P(q_t = S_i | q_{t-1} = S_j) : \text{state transition matrix} \]

\[ B = (b_{ij}) = P(v_k | S_j) : \text{observation symbol probability distribution} \]

*Notional Example

\[ A = \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.1 & 0.1 & 0.3 \end{pmatrix} \]

\[ B = \begin{pmatrix} 0.2 & 0.3 & 0.3 & 0.1 \\ 0.1 & 0.3 & 0.4 & 0.5 \end{pmatrix} \]

Outline

“Artificial Intelligence (AI) Methods” for Modeling and Assessing Collaborative Distance Learning

✔ Collaborative Distance Learning

✔ AI Methods: Intro to Hidden Markov Models

■ HMMs for analyzing online collaborative learning

■ HMMs and Neural Networks for assessing student adoption of problem solving strategies

■ Future project ideas
Collaborative Object Modeling Environment (COMET)

The COMET Chat Area

Logging the Interaction

A. Soller, 19 May 2006

Experimental Approach

1. Gather examples of student knowledge sharing
   - 12 groups of 3 students each
   - Each student received a different knowledge element
   - All 3 knowledge elements were needed to solve the problem
   - Students needed to share and discuss their knowledge elements
2. Tag examples: effective or not?
   - Pre-to-Post test improvement
3. Train HMMs to differentiate between effective & ineffective sequences
4. Re-train system to identify recurring types of knowledge sharing breakdowns
5. Use output to inform instructional module

A. Soller, 19 May 2006
Individual Knowledge Element Example

Attributes common to a group of subclasses should be attached to the superclass. This allows them to be shared by each subclass. Each subclass is said to inherit the features of its superclass. For example, *Cat* inherits the attributes name and color from *Pet*.

Example Training Sequence

<table>
<thead>
<tr>
<th>Student</th>
<th>Subskill</th>
<th>Attribute</th>
<th>Text Chat</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Request</td>
<td>Opinion</td>
<td><em>Do you think we need a discriminator for the car ownership?</em></td>
</tr>
<tr>
<td>C</td>
<td>Discuss</td>
<td>Doubt</td>
<td><em>I'm not so sure</em></td>
</tr>
<tr>
<td>B</td>
<td>Request</td>
<td>Elaboration</td>
<td><em>Can you tell me more about what a discriminator is?</em></td>
</tr>
<tr>
<td>C</td>
<td>Discuss</td>
<td>Agree</td>
<td><em>Yes, I agree because I myself am not so sure as to what its function is</em></td>
</tr>
</tbody>
</table>
| A       | Inform   | Explain   | *Let me explain it this way - A car can be owned by a person, a company or a bank. I think ownership type is the discriminator.*

(Note: Student "A" is always the knowledge sharer)
Training the HMMs

Actual HMM Training Sequence
A-Request-Opinion
C-Discuss-Doubt
B-Request-Elaboration
C-Discuss-Agree
A-Inform-Explain

14% A-Request-Info
12% C-Inform-Explain

12% C-Discuss-Doubt
10% C-Inform-Explain

12% B-Acknowledge-Accept
12% C-Motivate-Encourage
10% B-Discuss-Agree

Assessment by “Supervised” HMMs

Training
Effective knowledge sharing sequences

Ineffective knowledge sharing sequences

Testing
Knowledge sharing sequence

Probability that sequence is effective

Probability that sequence is ineffective
Results of Assessing Student Knowledge Sharing

- 29 examples of student knowledge sharing
  - 10 effective, 19 ineffective
  - Length of sequences ranged from 5-49 contributions
- HMMs achieved **74% accuracy** in distinguishing effective from ineffective knowledge sharing
  - Modified 58- fold cross validation, Baseline: 50%

*Hidden Markov Modeling is a useful tool for assessing sequences of knowledge sharing conversation & problem solving actions.*

Application and Potential Impact

*The students are having trouble understanding the concept of generalization*

Why?

- Instructional Module
  - Educational/Group Learning Theory applied for remediation
- Knowledge Sharing Analyzer
  - HMMs assess sequences of student activity
- Communication Interface & Shared Activity Workspace
- Dialog, Actions
- Analysis & Assessment Engine
  - (Running in Background)
Analysis of Knowledge Sharing Interaction

- Why are students having trouble?
  - Find & classify sequences with similar types of knowledge sharing breakdowns (or success)
  - Find closest match to current sequence

\[
\begin{align*}
M_1, & \ldots, M_i, \ldots, M_N \\
S_1 & \\
\vdots & \\
S_j & \\
\vdots & \\
S_N &
\end{align*}
\]

\[\loglik_j = \log L(S_j | M_i), \quad 1 \leq i, j \leq N\]

Columns describe the likelihoods of the sequences, given model \( M_i \).
So similar models should have similar column vectors.

HMM Clustering Procedure

- Multidimensional Scaling
  - To find distances between likelihood vectors
  \[L(S_j | M_i)\]
- ISODATA self organizing clustering routine
  - To find groups within threshold \( \theta \)

Results
- 4 ineffective groups
- 3 effective groups
Group C₁: Episodes: L8, L10, L12

S: Suggest KE
R1: Doubt that it is needed
S: Explain KE unsatisfactorily
R1: Doubt again
S: Request help from R2

S: Suggest KE
R1: Doubt suggestion
R1 & R2: Request elaboration of KE
S: Explain KE unsatisfactorily
(no further discussion of KE)

Soller, 19 May 2006

60-70% Coverage
Accuracy > 0.92

Group C₁: Episodes: L8, L10, L12

S: Suggest KE
R1: Doubt suggestion
R1 & R2: Request elaboration of KE
S: Explain KE unsatisfactorily
(no further discussion of KE)
**Application and Potential Impact**

"The students are having trouble understanding the concept of generalization..."

Mary is doubting Bob. Perhaps Bob could further justify his argument.

---

**Group C₂:** Episodes: L8, L10, L12

60-70% Coverage

Accuracy > 0.92
Summarized “Machine Learned” Knowledge Sharing Examples

Ineffective Knowledge Sharing
1. Sharer proposes KE
2. Sharer explains or gives instructions for action
3. Receiver acknowledges or requests confirmation
4. Receiver requests further clarification
5. Sharer provides further clarification

Effective Knowledge Sharing
1. Receiver requests information about KE
2. Sharer provides explanation
3. Receiver agrees

Lessons Learned

Hidden Markov Modeling is a useful tool for assessing sequences of knowledge sharing conversation & problem solving actions.

HMMs can be combined with other methods for assessing
How students successfully share knowledge or
Why students experience knowledge sharing breakdowns

A. Soller, 19 May 2006
Outline

“Artificial Intelligence (AI) Methods” for Modeling and Assessing “Collaborative Distance Learning”

- Collaborative Distance Learning
- AI Methods: Intro to Hidden Markov Models
- HMMs for analyzing online collaborative learning
  - HMMs and Neural Networks for assessing student adoption of problem solving strategies
  - Future project ideas

Interactive Multi-Media Exercises (IMMEX)

- IMMEX is an interactive, web-based, scientific problem solving environment
- Students learn to construct hypotheses, evaluate evidence, & draw inferences
- 100+ problem sets spanning middle school through medical school
- The IMMEX system models, assesses, and reports student progress in real-time.
- Neural Networks and HMMs analyze student problem solving strategies & predict future strategies
Layered Assessment in IMMEX

**Student Ability Estimates**
- Case item difficulties are determined by IRT analysis of 28,878 student performances.
- Each student is evaluated against this model.

**Problem Solving Strategy Models**
- Self-organizing Neural Networks cluster similar student performances in topologic maps.
- Each node represents a different problem solving strategy.

**Learning Progress Predictions**
- Sequences of Neural Network Nodes (from Layer 2) represent strategy changes over time.
- Nodes are sequenced stochastically to predict future learning.
An earthquake just hit your school

An unmarked container is damaged and the contents are spilling out.

Can you identify the chemical that was spilled so that you can dispose of it properly before it becomes a hazard to the school?
Assessing Student Adoption of Problem Solving Strategies

Students solve problems by applying a variety of strategies

Students shift their problem solving strategies over time

Neural Networks are used to identify a student's problem solving strategy for a given problem set

Hidden Markov Models are used to model students' shifting of problem solving strategies, and predict future learning

Defining Problem Solving Strategies with Neural Networks (1/2)
Defining Problem Solving Strategies with Neural Networks (2/2)

Item selection frequencies for 36 nodes trained with 5,284 student performances

Tracking Learning Progress

Student | Performance | Node Sequence
--- | --- | ---
1 | ![Performance graphs](image1.png) | [12, 18, 24, 20]
2 | ![Performance graphs](image2.png) | [5, 24, 6, 18]
3 | ![Performance graphs](image3.png) | [4, 22, 33, 33]
4 | ![Performance graphs](image4.png) | [32, 33, 28, 33]
Soller

Training HMMs to Represent Learning Progress Models

- Search background material extensively
- Strategically run tests

Tracking Learning Progress

- Search background material extensively
- Run many tests

A. Soller, 19 May 2006
Predicting Learning Trajectories

Collaboration Improves Performance & Accelerates Strategy Adoption

Average Solve Rate:
51% Individuals
63% Face-to-face groups
68% Online distributed groups
The Structured Interaction Model
(Alessandra Giordani, Ph.D. Candidate, University of Trento, Italy)

<table>
<thead>
<tr>
<th>Propose</th>
<th>Discuss</th>
<th>Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Let's try...</td>
<td>7. The test showed...</td>
<td>13. So far we know...</td>
</tr>
<tr>
<td>2. Why...?</td>
<td>8. What does that mean...?</td>
<td>14. We can eliminate...</td>
</tr>
<tr>
<td>3. We should...</td>
<td>9. Can you explain...?</td>
<td>15. If...</td>
</tr>
<tr>
<td>4. What do you think...?</td>
<td>10. It means...</td>
<td>16. Then...</td>
</tr>
<tr>
<td>5. Because...</td>
<td>11. Do you think...?</td>
<td>17. Do you know...?</td>
</tr>
<tr>
<td>6. I think (Free text proposal)</td>
<td>12. I think (Free text discussion)</td>
<td>18. I think (Free text review)</td>
</tr>
</tbody>
</table>

Problem solving phases and sentence openers

- **Propose**
  - Let’s try the inventory
  - We should click on the view the inventory sheet to see the message
  - Do you want to spend these points?
  - View Inventory
  - Let’s try either a pH test or silver nitrate test to identify the pH or whether we have a chloride.

- **Discuss**
  - We should click on the view the inventory sheet to see the message
  - Do you want to spend these points?
  - Yes
  - The test showed we have a sodium cation so we have carbonate, chloride, hydroxide, nitrate, or sulfate anions.

- **Review**
  - OK
  - It means... |


Analysis of Collaborative Learning Dialog

- Four pairs of students collaboratively performed 15 IMMEX distance learning cases
- More proposals occur during the early framing stages of problem solving (88% cases)
- As the students converged upon a solution, there was proportionally more discussion (69% cases)
- In 94% of the chat logs, the proposal rates decreased in the second half of the dialog, and the discussion increased (from 25% to 64%)
Summary & Future Project Ideas

- What hybrid combinations of machine learning & mixture models help to assess which aspects of learning and collaboration?
  - Project Idea: Test & Compare performance of Neural Nets, Decision Trees, HMMs, Bayesian models

- To what extent can predictive properties of these models enable focused experimentation of problem-solving interventions?
  - Project Idea: Explore effect of selecting intervention based on intelligent prediction vs. traditional methods
Selected References

(available at http://www.cscl-research.com or upon request to asoller@ida.org)


Thank You!

http://www.cscl-research.com

asoller@ida.org