UDC 004.78:025.4.036

# Machine learning models of problem space navigation: the influence of gender

Ron Stevens, Amy Soller

IMMEX Project, UCLA, 5601 W. Slauson Ave, #255, Culver City, CA. 90230

immex_ron@hotmail.com

asoller@ida.org

**Abstract.** We have developed models of how problem spaces are navigated as male and female secondary school, university, and medical students engage in repetitive complex problem solving. The strategies that students used when solving problem-solving simulations were first classified with self-organizing artificial neural networks resulting in problem solving strategy maps. Next, learning trajectories were developed from sequences of performances by Hidden Markov Modeling that stochastically described students' progress in understanding different domains. Across middle school to medical school there were few gender differences in the proportion of cases solved; however, six of the seven problem sets showed significant gender differences in both the strategies used (ANN classifications) as well as the in the HMM models of progress. These results were extended through a detailed analysis of one problem set. For this high school / university problem set, gender differences were seen in how the problems were encoded, consolidated and retrieved. These studies suggest that strategic problem solving differences are common across gender, and would be masked by simply looking at the outcome of the problem solving event.

## 1. Introduction

Research documenting how mammals represent the environment at the behavioral, cellular and molecular levels (Hasselmo et al, 1996; Eichenbaum, 2000) is suggesting a world-centered representation of the environment organized mainly in the hippocampus. Here, whenever an animal is in a particular location in a particular environment specific place cells become activated. The activities of these place cells are stable across days and weeks in constant surroundings, but change in association with environmental changes (Arleo & Gerstner, 2000; Nakazawa et al, 2004). These 'remappings' suggest that the hippocampus learns and retrieves distinct maps for distinct contexts, with attractors within each specific map being reactivated (presumably while drawing on other neural contributions such as the neocortex and parahippocampus). More recent modeling studies suggest that such attractors, and the remapping phenomena may be a manifestation of a more general capacity to associate arbitrary conjunctions of fragments of experiences and thus a more general feature of memory (McClelland & Goddard, 1996; Willis et al, 2005).

Continuous navigation in familiar and unfamiliar environments is also one of the few behaviors for which reliable gender-specific cognitive performance differences are well documented in humans (Tlauka et al, 2005; Gron et al, 2000). For instance, when knowledge is acquired through real-world navigation, males typically outperform females on subsequent tests of spatial ability. These results have been extended to navigation through computer-simulated spaces and are being generalized to suggest that women are more inclined to use landmarks for navigation, while men may rely more on a sense of geometrical direction.

Such navigation cues, however, may extend past the tangible world to a more abstract form of navigation, the navigation of problem spaces. For instance, the differential spatial propensities involved in real-world navigation have also been observed in items in the SAT-M and the GRE-Q item pools where, as a group, males outperform females on items requiring spatial skills, shortcuts, or multiple solutions, while females tend to outperform males on problem requiring verbal skills or mastery of classroom-based content (Gallagher et al, 2000).

Gender-based strategic differences in solving problems may, however, be more common than previously thought as they may go un-noticed as a result of little or no differences in the immediate performance criteria of the problem, which, for most cases, is whether or not the item was answered correctly. An example of the possible dichotomy between performance and strategic measures was suggested in a longitudinal study of math learning by Fennema et al (1995) who studied teachers and their students as they progressed from Grades 1 through 3. While they found no gender differences in the solved rate of problems performed during the study, there were strong and consistent differences in the strategies used to solve problems with boys favoring the use of invented rather than procedural algorithms for solving problems. These gender trends regarding problem solving are sometimes generalized as males being more risky and females more conservative in their strategic approaches. However, the gender correlates of 'risky' and 'conservative' have not been widely expanded on experimentally.

We are attempting, therefore, to recast some of these findings into a schema for problem solving which may lead to links, at a cognitive level, between directional and wayfinding tasks in the real world, and the abstract representation of the environment during complex problem solving. For instance, is a 'conservative' problem-solving approach similar to setting multiple landmarks during navigation, and is a 'risky' approach similar to finding spatial shortcuts or detours? If so, can gender differences be observed both during encoding (the first exploration of a problem space) and consolidation (when the task is repeated) as experience is gained within a problem solving environment? Continuing the theme of neural correlates, is it also possible to begin to recognize pattern separation events and pattern generalization events within a problem solving system, similar to those that have been shown to exist in navigation and memory encodings and retrievals?

Most studies of gender differences in problem solving have involved the domain of mathematics and it is not clear how common the findings are across different science domains and grade levels. It is also not clear if the separation of problem solving performance and strategy observed by Fennema et al (1998), and the associated gender effects, represents an unusual finding, or one more generally representative of math and science learning.

The above questions suggest that an important next step would be to model student strategies across different domains, at various levels of detail and also with regard to gender as problem solving experience develops, to not only provide evidence of students' ability to solve problems, but also of their changing strategic task understanding.

We have begun to address these questions with an online problem-solving system, collectively called IMMEX (Interactive Multi-Media Exercises), (Stevens et al, 1999; Underdahl et al, 2001). IMMEX problem solving follows the hypothetical-deductive learning model of scientific inquiry (Lawson, 1995) where students need to frame a problem from a descriptive scenario, judge what information is relevant, plan a search strategy, gather information, and eventually reach a decision that demonstrates understanding (http://www.immex.ucla.edu).

In this study, we describe a process to model strategic development using a sample problem set termed *Hazmat* that contains 38 different cases and provides evidence of a student's ability to conduct qualitative chemical analyses (Figure 1). The problem begins with a multimedia presentation explaining that an earthquake caused a chemical spill in the stockroom, and the student's challenge is to identify the chemical. The problem space contains 22 menu items for accessing a Library of terms, the Stockroom Inventory, or for performing Physical or Chemical Testing. When the student selects a menu item, she verifies the test requested and is then shown a multimedia presentation of the test results (e.g. a precipitate forms in the liquid, or the light bulb switches on suggesting an electrolytic compound). When students feel they have gathered adequate information to identify the unknown they can attempt to solve the problem. The IMMEX database collects timestamps of each student selection.

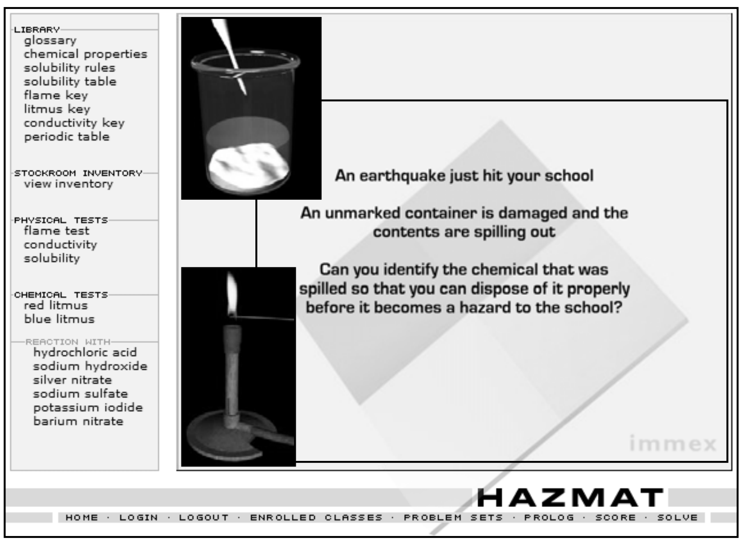


**Fig. 1.** *HAZMAT* This composite screen shot of *Hazmat* illustrates the challenge to the student, and shows the menu items on the left side of the screen. Also shown are two of the test items available. The item in the upper left corner shows the result of a precipitation reaction and the frame at the lower left is the result of flame testing the unknown.

Over 100 IMMEX problem sets have been created that span middle school to medical school, reflect disciplinary learning goals, and meet state and national curriculum objectives and learning standards. We have selected data from seven of these problem sets for these studies

To model students' performance and progress we have developed layered analytic processes for determining how strategies are constructed, modified and retained as students learn to solve problems like *Hazmat*. These layers operate as background processes and can generate most performance measures in real-time (Stevens & Soller, 2005).

***Layer 1. Item Response Theory Estimates of Student Ability.*** To ensure that students gain adequate experience, the *Hazmat* problem set contains 38 cases with a variety of acids, bases and compounds giving either a positive or negative result when flame tested. As students perform multiple cases that vary in difficulty, refined estimates of student ability are obtained by IRT analysis by relating characteristics of items and individuals to the probability of solving a given case (Linacre, 2004). As expected, the flame test negative compounds are more difficult for students because both the anion and cation have to be identified by running additional chemical tests. Overall, the problem set presents an appropriate range of difficulties to provide reliable estimates of student ability.

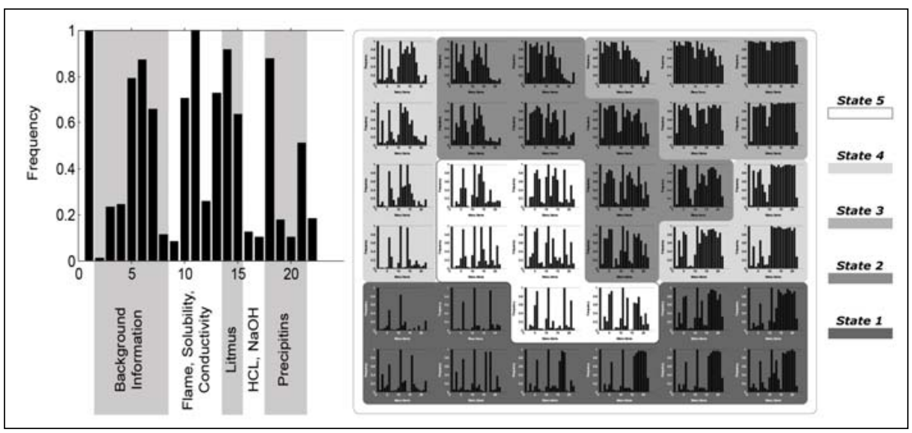


**Fig. 2.** Sample Neural Network Nodal Analysis. **a.)** The selection frequency of each action (identified by the labels) is plotted for the performances at node 15, and helps characterize the performances clustered at this node and for relating them to performances at neighboring nodes. **b.)** This figure shows the item selection frequencies for all 36 nodes, and maps them to HMM states.

***Layer 2. Artificial Neural Network (ANN) Classification of Strategies.*** While useful for ranking the students by the effectiveness of their problem solving, IRT does not provide strategic measures of this problem solving. Here, we use ANN analysis. As students navigate the problem spaces, the IMMEX database collects timestamps of each student selection. The most common student approaches (i.e. strategies) for solving *Hazmat* are identified with competitive, self-organizing artificial neural networks (Kohonen, 2001; Stevens & Najafi, 1993, Stevens et al, 1996) using these time-stamped actions as the input data. The result is a topological ordering of the neural network nodes according to the structure of the data where geometric distance becomes a metaphor for strategic similarity.

Often we use a 36-node neural network and the classification results are visualized by histograms showing the frequency of items selected for student performances classified at that node (Figure 2 A). Strategies so defined consist of actions that are always selected for performances at that node (i.e. with a frequency of 1) as well as ones ordered variably. Figure 2 B is a composite ANN nodal map that shows the topology of performances generated during the self-organizing training process. Each of the 36 matrix graphs represents one ANN node where similar student's problem solving performances have become competitively clustered. As the neural network was trained with vectors representing selected student actions, it is not surprising that a topology developed based on the quantity of items. For instance, the upper right of the map (nodes 6, 12) represents strategies where a large number of tests were ordered, whereas the lower left contains strategies where few tests were ordered. Once ANN's are trained and the strategies represented by each node defined, new performances can be tested on the trained neural network and the node (strategy) that best matches the new performance can be identified and reported.

*Layer 3. Hidden Markov Model (HMM) Strategic Progress Models.* On their own, artificial neural network analyses provide point-in-time snapshots of students' problem solving. Any particular strategy, however, may have a different meaning at a different point in a learning trajectory. More complete models of student learning should also account for the changes of student's strategies with practice. Our approach here is to have students perform multiple cases in the 38-case *Hazmat* problem set, and classify each performance with the trained ANN. Predictive models of student learning trajectories are then developed from sequences of these strategies with HMM (Rabiner, 1989; Murphy, 2004).

The critical components of such an analysis are shown in Figure 3 where students solved 7 *Hazmat* cases. One level (stacked bar charts) shows the distribution of the 5 HMM states across the 7 performances.

On the first case, when students are framing the problem space, the two most frequent states are States 1 and 3. Moving up an analytical layer from HMM states to ANN nodal strategies (the 6 x 6 histogram matrices) shows that State 3 represents strategies where students ordered all tests, and State 1 where there was limited test selection. Consistent with the state transitions in the upper right of Figure 3, with experience students transited from State 3 (and to some extent State 1), through State 2 and into States 4 and 5, the more effective states. By the fifth performance the State distributions stabilized after which time students tend not to switch their strategies, even when they were ineffective.

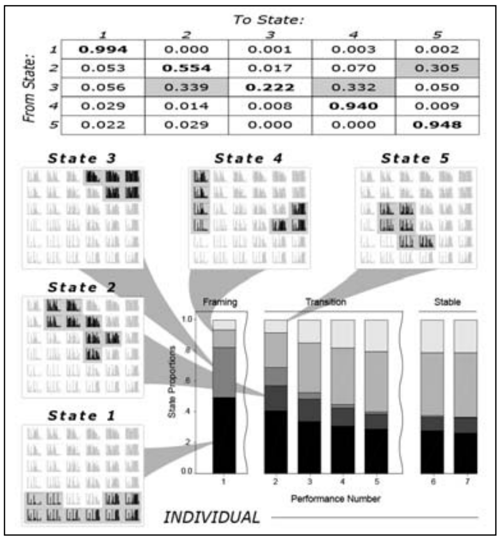


**Fig. 3.** Modeling Individual Learning Trajectories. This figure illustrates the strategic changes as individual students gain experience in *Hazmat* problem solving. Each stacked bar shows the distribution of HMM states for the students (N=1790) after a series (1-7) of performances. These states are also mapped back to the 6 x 6 matrices which represent 36 different strategy groups identified by self organizing ANN. The highlighted boxes in each neural network map indicate which strategies are most frequently associated with each state. From the values showing high cyclic probabilities along the diagonal of the HMM transition matrix (upper right), States 1, 4, and 5 appear stable, suggesting once adopted, they are continually used. In contrast, students adopting State 2 and 3 strategies are more likely to adopt other strategies (gray boxes).

*Hazmat Gender Influences* – If indeed females are more prone to thoroughly explore problem spaces to establish landmarks, then within the context of IMMEX problem solving this could be differentially reflected in the quantity of tests chosen. Consistent with this hypothesis, on the first case of the problem set, there was a significant difference in state usage with males preferring State 1 approaches and females preferring State 3 approaches (Pearson $\chi2$ = 13.54, P=0.004). State 1 consists of Nodes 25, 26, and 29-36 which represent performances often of the limited data type. State 3 is just the opposite where many of the available background and test items are selected. During subsequent cases, there was a steady reduction in States 1 and 3, a transient appearance of State 2 performances and then the emergence of States 4 and 5 performances. Males tended to use State 4 strategies while more females progressed to State 5 strategies. (Pearson $\chi2$ = 31.2, p<0.000).

***Hazmat Strategy Stabilization and Persistence*** - We next explored the stability of student's strategies and the influences of gender. These studies were performed with a smaller set of advanced placement chemistry students that included 6 classrooms of the same teacher. The studies were conduced with 3 of these classes in Spring 2004 and replicated in 3 additional classes in Spring 2005; the total number of students was 182, and the total number of performances was1932. In a standard classroom environment students first performed 5-6 *Hazmat* problems to refine and stabilize their strategies. Then 15 weeks (Spring 2004) or 6 weeks (Spring 2005) students were asked to solve additional *Hazmat* cases. The overall solution frequency slightly favored females (F= 65.5 % solved, Males 60.1% solved, Pearson $\chi2$ = 6.3, p=.043). The performances were then separated by gender and the state distributions were re-plotted. Overall, males had higher than expected numbers of States 1 and 4 performances, and females had more State 5 performances (Pearson $\chi2$ = 34.9, p<.000).

At the end of the required first-set of performances (# 1-5), the proportions of the five strategy states and the solution frequencies had stabilized for both males and females. However, as shown in Figure 4, while both male and female students appeared to have stabilized their strategic approaches by the fifth performance, the state distributions were significantly different, with females preferring the approaches represented by State 5 and males preferring State 4 approaches.

***A Survey of Problem Solving Gender Effects Across Grade Levels and Domains*** - We next expanded these studies to survey both the solution frequencies as well as the strategic profiles across gender, grade levels, and scientific domains. This survey included examples from middle and high school, university as well as medical school students, and we selected problem sets that have been used by multiple classes and schools and where over 3,000 performances had been collected. As shown in Table 1, for six of the seven problem sets there were no significant gender differences in the solution frequencies. At the ANN nodal strategic level, with the exception of the medical school microbiology and the university molecular biology problem sets, there were however, significant gender differences in the strategies used by students to solve the cases.

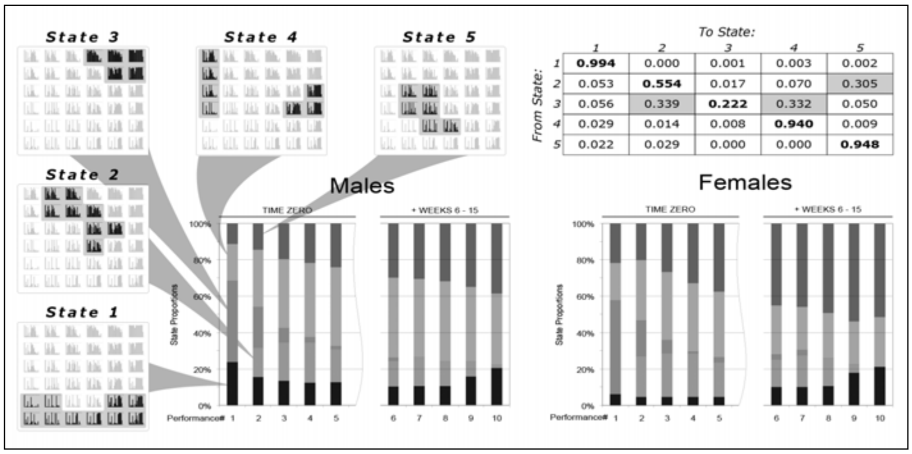


**Fig. 4.** Development and Persistence of Male and Female Learning Trajectories. This figure illustrates the strategic changes of males and females in six Advanced Placement Chemistry courses taught by the same teacher. For reference, the boxed histograms map the most likely ANN node to the different states, and the matrix in the upper right hand corner shows the training HMM transition matrix illustrating the likelihood of moving across states.

**Table 1.**

<table>
<tr><td colspan="2">Problem Set</td><td colspan="2">Description</td><td>Grade Level</td></tr>
<tr><td colspan="2">Microbiology</td><td colspan="2">Use your knowledge of microbiology to diagnose a patient’s symptoms.</td><td>Medical School</td></tr>
<tr><td colspan="2">Solve Rate</td><td colspan="2">Strategy Differences</td><td>Description,</td></tr>
<tr><td>Male</td><td>75%</td><td rowspan="2">Node vs. Gender</td><td rowspan="2">State vs. Gender</td><td rowspan="3">High case specificity. Males – more extensive testing.</td></tr>
<tr><td>Female</td><td>75%</td></tr>
<tr><td>$\chi^2$(df)=,p=</td><td>(2).373,.803</td><td>(34)45.7,.106</td><td>(4)31.2,.000</td></tr>
<tr><td colspan="2">Problem Set</td><td colspan="2">Description</td><td>Grade Level</td></tr>
<tr><td colspan="2">lac Operon (Molec. Biology)</td><td colspan="2">From an isolated colony identify the exact mutation using the resources available.</td><td>University</td></tr>
</table>

| Solve Rate | | Strategy Differences | | Description, |
|---|---|---|---|---|
| Male | 79% | Node vs. Gender | State vs. Gender | None |
| Female | 75.4% | | | |
| $\chi^2$(df)=,p= | (1)3.8,.055 | (4)3.26,.515 | (34)38.9,.257 | |

| Problem Set | Description | Grade Level |
|---|---|---|
| **Hazmat** (Qual. Analysis) | Identify the chemical that has spilled so that you can help dispose of it properly. | University/High School |

| Solve Rate | | Strategy Differences | | Description, |
|---|---|---|---|---|
| Male | 53% | Node vs. Gender | State vs. Gender | Males appear to partition the problem space while females develop more generalizable strategies. |
| Female | 50.6% | | | |
| $\chi^2$(df)=,p= | (1)2.8,.092 | (34)52.7,<.000 | (4)31.2,<.000 | |

| Problem Set | Description | Grade Level |
|---|---|---|
| **Roger Rabbit (Forensic Sci.)** | Determine which students are responsible for poisoning the class pet. | Middle/High School |

| Solve Rate | | Strategy Differences | | Description, |
|---|---|---|---|---|
| Male | 32% | Node vs. Gender | State vs. Gender | Males using more non-science and guessing strategies. Females more extensive and test oriented |
| Female | 36% | | | |
| $\chi^2$(df)=,p= | (2)5.61,.06 | (34)112.36,.000 | (4)57.48,.000 | |

| Problem Set | Description | Grade Level |
|---|---|---|
| **Duck Run** (Chemistry) | A mysterious element has been dumped into the local duck pond. Identify the element. | Middle School |

| Solve Rate | | Strategy Differences | | Description, |
|---|---|---|---|---|
| Male | 72% | Node vs. Gender | State vs. Gender | Females more extensive navigation than males. Males more guessing. |
| Female | 71% | | | |
| $\chi^2$(df)=,p= | (1).385,.535 | (34)107,<.000 | (4)43,<.000 | |

| Problem Set | Description | Grade Level |
|---|---|---|
| **RXN** | Complete a data table by solving for the unknown chemical equation. | Middle School |

| Solve Rate | | Strategy Differences | | Description |
|---|---|---|---|---|
| Male | 45% | Node vs. Gender | State vs. Gender | Females use data-oriented problem solving approaches, and males relied more on guessing. |
| Female | 51% | | | |
| $\chi^2$(df)=,p= | (1)20.4,.000 | (35)105.7,<.000 | (4)60.0,<.000 | |

| Problem Set | Description | Grade Level |
|---|---|---|
| **Road Trip** | Calculate distance, travel time, and route. | Middle School |

| Solve Rate | | Strategy Differences | | Description |
|---|---|---|---|---|
| Male | 55% | Node vs. Gender | State vs. Gender | Both groups are doing extensive exploration with males exhibiting more shortcuts and guessing. |
| Female | 53% | | | |
| $\chi^2$(df)=,p= | .605(1),.437 | (35)267,<.000 | (4)99, <.000 | |

## 2. Discussion

The influence of gender on strategic learning and progress during complex problem solving has been modeled using ANN and HMM machine modeling technologies. The resulting data indicate that while the solution frequency for each of the seven problem sets analyzed was similar for males and females, the performance models as measured by ANN analysis and the progress models as measured by HMM were consistently different. These results suggest that differences in the ways that males and females navigate problem spaces may be a common occurrence. These findings are contingent, of course, on the validity of the tasks as well as the performance and strategic models developed from the student data.

In these studies we have focused on validating one representative problem set, *Hazmat*, where over 34,000 performances have been recorded by high school and university students. This problem set was created along the frameworks we have published previously (Stevens & Palacio-Cayetano, 2003). The *Hazmat* task has face validity in that it covers much of the spectrum of qualitative analysis with the 38 parallel cases that include acids, bases, and flame test positive and negative compounds. The tasks also have construct validity in that cases are of different difficulties by Item Response Theory analysis (Linacre, 2004), and these differences correlate with the nature of the compounds (e.g. flame test positive compounds are easier than flame test negative compounds).

The next validation step addresses the quality of the collected data. IMMEX simulations require that students re-confirm each test ordered (for a cost), resulting in a series of deliberate actions. These actions also have cognitive correlates as concurrent verbal protocol analysis has indicated that ~90% of the utterances by students can be categorized into explicit cognitive or metacognitive categories (Chung et al, 2001).

In the first modeling step the most common strategies used by students were grouped by unsupervised ANN analysis and the resulting classifications showed a topology ranging from those where very few tests were ordered, to those where every test was selected, which makes sense given the nature of the input data (i.e. deliberate student actions). The HMM progress models are somewhat more difficult to validate given the hidden nature of the model. One important consideration would be the dynamics of the state transitions as reflected in the transition matrix derived from the modeling process. Here theories of practice and cognition (Ericsson, 2004) would predict that students would change strategies with practice and eventually stabilize with preferred approaches much as we have shown in Figures 3 and 4. Similarly, the general overall shift in states from those representing extensive exploration to more refined test selection mirrors the data reduction effects observed previously with practice (Haider & Frensch, 1996). For instance, most students approached the first *Hazmat* case by selecting either an extensive (State 3), or limited/guessing (State 1) amount of information. The State 3 approaches would be appropriate for novices on the first case as they define the boundaries of the problem space. Persisting with these strategies, however, would indicate a lack of understanding and progress.

As students gain experience, their strategies should change (Ericsson, 2004). Background information that was needed earlier may no longer be needed, and students should begin to develop their own preferred approaches based on knowledge, experience, motivation, and prior experiences. The main transition states are States 2 and 3. When students transition out of State 3, this suggests that they are learning, and the transition matrix shows that these students are likely to switch to States 2 or 4 increasing their likelihood of solving the case from 27% to 40%. The main difference between States 2 and 4 is that there are both test and background information being accessed in the State 2 approaches whereas State 4 includes primarily data driven approaches.

The states that students stabilize with presumably reflect their level of competence as well as the approaches they are comfortable with, and are the ones that would most often be recognized by teachers. For *Hazmat* the stabilized states were represented by States 1, 4 and 5. State 4 is interesting in several regards. First, it differs from the other states in that the strategies it represents are located at distant points on the ANN topology map, whereas the nodes comprising the other states are contiguous. The State 4 strategies represented by the left hand of the topology map are appropriate for the set of cases in *Hazmat* that involve flame test positive compounds, whereas those strategies on the right are more appropriate for flame test negative compounds (where more extensive testing for both the anion and cation are required). This suggests that students using State 4 strategic approaches may have mentally partitioned the *Hazmat* problem space into two groups of strategies, depending on whether the initial flame test is positive. At the cognitive level, this is suggestive of pattern separation events postulated to occur in the hippocampus.

State 5 also contains complex strategies which from the transition matrix emerge from State 2 strategies by a further reduction in the use of background resources. State 5 approaches appear later in problem solving sequences, have the highest solution frequencies and are approaches that work well with both flame test positive and negative compounds. In this regard they may represent the outcome of a pattern consolidation process. They are also strategies used more frequently by females than by males.

What can these studies tell us about gender and problem solving? First, they suggest that the way complex problem spaces are navigated and encoded have a significant gender component which can be observed from middle school to early university years. It may be important that the two problem sets not showing these differences were both created for and performed by more advanced university/medical students. The lack of gender differences may reflect the combination of a highly selected group of students as well as the nature of the problems created for such students. For instance, in the medical school environment, the cases are often framed (encoded) for the students by the organ system(s) involved, the patient's age, the acute vs. chronic nature of the illness, etc. The more general problem solving involved with earlier grade level problem sets may allow more, and a greater variety of initial representations to be encoded, consolidated and eventually modeled.

However, the gender trends are significant and do suggest that gender may be used as a tool to further explore aspects of learning and problem solving. For instance, from the strategic learning trajectories in Figures 3 and 4 both males and females appear to stabilize their strategies within the period of an hour, and these persist for several months, both properties reflective of the fixation of cellular memory component of the memory consolidation processes (McGaugh, 1966). Nevertheless, while males and females share temporal consolidation characterizations, the stabilization with different state distributions suggests other mechanisms are helping to shape the quality of the strategy adopted.

Second, these studies help suggest a new criterion for designing complex instructional materials. While high stakes testing efforts have attempted to neutralize gender differences in outcome measures by removal of items that strategically favor males or females (Gallagher, 1994), our tasks appear to have few gender differences in outcomes, yet allow significant gender-related approaches to occur.

Third, this separation of performance outcome (solving the problem) and strategic approach itself warrants additional study. How related are these two constructs and are they differentially affected by the factors such as task assignment, classroom environment, and prior instruction?

These studies have not however directly addressed the question posed in the introduction as to the relation of navigation in the real world and the more abstract navigation of complex problem spaces. Here, it may be useful to look for the establishment of domain-related and gender-influenced landmarks as problem solving experience is gained across a domain such as qualitative chemistry. Once students have encoded preferred strategies on one problem set such as *Hazmat*, are some of the dominant landmarks (sequences of test selections) established selectively carried forward when more complicated problems in qualitative analysis (such as those involving redox reactions) are presented in a transfer task? Several such pairs of problem sets exist where adequate numbers of performances have been collected for such modeling, and others can be created for studying such effects.

Supported in part by grants from the National Science Foundation (ROLE 0231995, DUE 0126050, HRD-0429156).

## References

1. Arleo, A. and Gerstner, W. (2000). Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity. Biological Cybernetics, 83, 287-299.
2. Chung G. R. de Vries L. F. Cheak A. M. Stevens R. H. Bewley W. L. (2001). Cognitive process validation of an online problem solving assessment. Comput. Hum. Behav. 2002;18:669–684.
3. Corbett, A. T. & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 4, 253-278.
4. Eichenbaum, H. (2000). A Cortical-hippocampal system for declarative memory. Nature Reviews Neuroscience 1, 41 -50.
5. Ericsson, K.A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Academic Medicine, 79, S70-S81.
6. Fennema, E. Carpenter, T., Jacobs, V., Franke, M., and Levi, L. (1998). Gender differences in mathematical thinking. Educational Researcher, 27, 6-11.

7. Gallagher, A. M., De Lisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V. Morely, M. And Cahalan, C. (2000). Gender differences in advanced mathematical problem solving. Journal of Experimental Child Psychology. 75, 165-190.

8. Gallagher, A.M., and De Lisi, R. (1994). Gender differences in scholastic aptitude test mathematics problem solving among high ability students. Journal of Educational Psychology, 86(2), 204-214.

9. Grön, G., Wunderlich, A. P., Spitzer, M., Tomczak, R., and Riepe, M. W. (2000). Brain activation during human navigation: Gender-different neural networks as substrate of performance. Nature vol. 3(4); 404-408.

10. Guzowski J.F., Knierim J.J., and Moser E.I. (2004) Ensemble dynamics of hippocampal regions CA3 and CA1. Neuron. 18; 44(4): 581-4.

11. Haider, H., and Frensch, P.A. (1996). The role of information reduction in skill acquisition. Cognitive Psychology 30: 304-337

12. Hasselmo ME, Wyble BP, and Wallenstein GV. (1996). Encoding and retrieval of episodic memories: role of cholinergic and GABAergic modulation in the hippocampus. Hippocampus; 6:693 – 708.

13. Kohonen, T., 2001. Self Organizing Maps. 3rd extended edit. Springer, Berlin, Heidelberg, New York

14. Lajoie, S.P. (2003). Transitions and trajectories for studies of expertise. Educational Researcher, 32: 21-25.

15. Lawson, A.E. (1995). Science Teaching and the Development of Thinking. Wadsworth Publishing Company, Belmont, California

16. Linacre, J.M. (2004). WINSTEPS Rasch measurement computer program. Chicago. Winsteps.com

17. McClelland, J. L., and Goddard, N. H. (1996) Considerations arising from a complementary learning systems perspective on hippocampus and neocortex. Hippocampus 6:654-665.

18. McGaugh, J.L., (1966). Time-dependent Processes in Memory Storage. Science 153, 1351-1358.

19. Murphy, K. (2004). Hidden Markov Model (HMM) Toolbox for Matlab. Available online at: http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html.

20. Nakazawa, K, McHugh, T. J., Wilson, M. A., and Tonegawa, S. (2004). NMDA receptors, place cells and hippocampal spatial memory. Nature Reviews Neuroscience 5, 361 -372.

21. Olson, A., & Loucks-Horsley, S. (Eds). (2000). Inquiry and the National Science Education Standards: A guide for teaching and learning. Washington, DC: National Academy Press

22. Rabiner, L., (1989). A tutorial on Hidden Markov Models and selected applications in speech recognition. Proc. IEEE, 77: 257-286

23. Stevens, R., Wang, P., Lopo, A. (1996). Artificial neural networks can distinguish novice and expert strategies during complex problem solving. JAMIA vol. 3 Number 2 p 131-138

24. Stevens, R.H., and Najafi K. (1993). Artificial neural networks as adjuncts for assessing Medical students' problem-solving performances on computer-based simulations. Computers and Biomedical Research 26(2), 172-187

25. Stevens, R.H., Ikeda, J., Casillas, A., Palacio-Cayetano, J., and S. Clyman (1999). Artificial neural network-based performance assessments. Computers in Human Behavior, 15: 295-314

26. Tlauka, M., Brolese, A., Pomeroy, D., and Hobbs, W. (2005). Gender differences in spatial knowledge acquired through simulated exploration of a virtual shopping Centre. Journal of Environmental Psychology. 25; 111-118.

27. Underdahl, J., Palacio-Cayetano, J., and Stevens, R. (2001). Practice makes perfect: assessing and enhancing knowledge and problem-solving skills with IMMEX software. Learning and Leading with Technology. 28: 26-31

28. VanLehn, K., (1996). Cognitive Skill Acquisition. Annu. Rev. Psychol 47: 513-539

29. Wills, T. J., Lever, C. Cacucci, F., Burgess, N., and O'Keefe, J. (2005). Attractor dynamics in the hippocampal representation of the local environment. Science vol. 308 873-876.