JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

Making Sense of Students' Actions in an Open-Ended Virtual Laboratory Environment

Ya'akov (Kobi) Gal,[‡] Oriel Uzan,[‡] Robert Belford,[§] Michael Karabinos,[†] and David Yaron*,[†]

[†]Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

[‡]Department of Information Systems Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

[§]Department of Chemistry, University of Arkansas at Little Rock, Little Rock, Arkansas 72204, United States

Supporting Information

ABSTRACT: A process for analyzing log files collected from open-ended learning environments is developed and tested on a virtual lab problem involving reaction stoichiometry. The process utilizes a set of visualization tools that, by grouping student actions in a hierarchical manner, helps experts make sense of the linear list of student actions recorded in a raw log file. Such analysis of an initial set of log files is then used to develop a rule-based system that can automatically classify the problem-solving strategies being engaged in by the students. The strategies assigned by the resulting rule-based system compare well with strategy codes assigned by experts.

KEYWORDS: *First-Year Undergraduate/General, High School/Introductory Chemistry, Chemical Education Research, Computer-Based Learning, Inquiry-Based/Discovery Learning, Internet/Web-Based Learning, Stoichiometry*

FEATURE: Chemical Education Research

INTRODUCTION

A growing theme of modern chemical education is engaging students in a range of science practices that better reflect the ways chemists inquire about the world.[1−3] By learning chemical content while engaging in practices such as the design and interpretation of experiments, students may develop a deeper and more memorable conceptual understanding. Virtual laboratories have the potential to broaden access to such educational experiences. The ChemCollective virtual laboratory (vlab) studied here couples a general chemical simulation with a flexible user interface that gives students wide latitude in design of their experiments. The vlab is therefore a type of Exploratory Learning Environment (ELE), a class of open-ended and flexible educational software that allows students to build scientific models and examine properties of the models by running them and analyzing the results.[4−6] The flexible and open-ended nature of ELEs make for a rich educational environment for students. However, the open-ended character of the educational experience increases the need for student support. In both classroom and online environments, the provision of such support is hindered by the lack of information regarding what students have done in the online environment. This is reflective of a major challenge associated with using ELEs in educational settings, that of providing students with the right kind of help, feedback and other forms of scaffolding required to make learning efficient.[4] The complexity associated with providing scaffolding can be reduced by constraining the user interface such that students, at any given point, can select from only a few options.[7] However, such constraints may compromise the learning benefits of engaging in open-ended problem solving.[8−10] This paper advances a key technology needed to support student work within ELEs without imposing artificial constraints on their problem solving choices. This key technology is an automated system that can, given only the interactions of the student with the software, classify the student's problem-solving strategy.

Our goal is to take the raw interactions of the students with the software, stored as mouse clicks in a log file or streamed to a server online, and from this information infer the strategies the student is using to solve the chemistry problem. Our process begins with human analysis of log files, but uses the information gathered from this human analysis to develop a computer system that can automatically classify problem solving strategies. This process is developed and tested on a problem solving task, within the vlab, that involves identification of an unknown chemical reaction involving four chemical species. This problem solving task is sufficiently open-ended to elicit a number of distinct problem-solving strategies.

Our process is based on the following hierarchical grouping of the students' actions in the virtual lab, which is shown in Figure 1:

- **Actions** are the basic forms of interaction, such as pouring 10 mL of solution from one flask into another. A list of these actions is saved in a text log file, whose length typically varies from 20 to 100 actions.
- **Activities** are sets of actions that correspond to a laboratory procedure, such as using a pipet to transfer a solution between two vessels, mixing two solutions, or carrying out a titration. Activities correspond to common laboratory manipulations but are not meant to describe intent regarding problem solving.
- **Strategies** are sets of activities carried out with a specific intent regarding problem solving. An example strategy

DOI: 10.1021/ed500531a
J. Chem. Educ. XXXX, XXX, XXX−XXX

for the problem solving task considered here is mixing all pairs of chemical species to discover which react.
- **Solution paths** are sets of strategies that, taken together, allow one to solve a given problem. For the problem solving task considered here, a typical solution path involves two types of strategies, initial strategies that help identify the chemical species that react and additional strategies that help identify the full stoichiometry of the reaction.

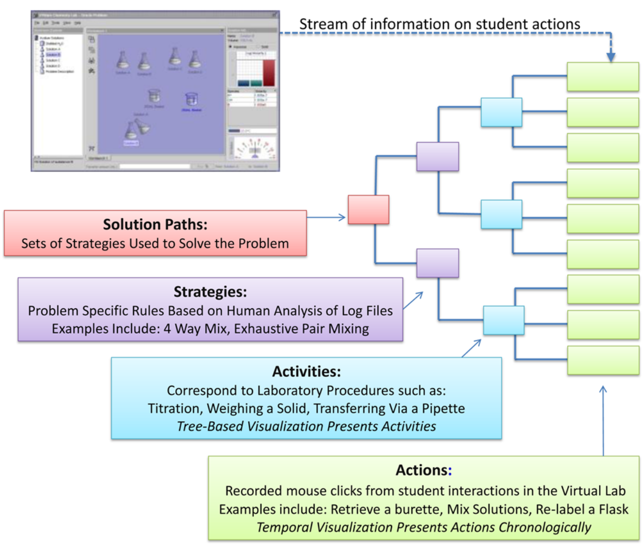

Figure 1. Schematic representation of the hierarchy used to categorize student problem solving in an open-ended virtual lab activity.

The log file consists of a stream of actions performed by the student and the goal of the analysis is to group these actions into the hierarchy of Figure 1. An automated means for grouping actions into activities has already been developed and is utilized in the current work. This level of the hierarchy captures student intent regarding only common laboratory manipulations and so the process used to form the groups applies across multiple problem types. The current work extends the analysis to the upper levels of the hierarchy, where the groupings are meant to capture student intent regarding problem solving. An approach to forming such groups that would apply across multiple problem types is difficult to envision. Instead, the approach discussed in the Methodology section is adopted in which domain experts develop a coding scheme for detecting the presence of problem-solving strategies and solutions paths in a log file. An automated means is then developed for grouping actions into the strategies of this coding scheme. The result is a computing system that can, by grouping actions into the hierarchy of Figure 1, analyze a log file to yield information regarding students' problem-solving.

To analyze a log file, the raw actions performed by the student are first grouped into activities. The resulting activities are then grouped into strategies, with the set of all strategies identified in the log file corresponding to the solution path. As discussed in the Methodology section below, the methodology used to group actions into activities differs from that used to group activities into strategies. These differences in methodology relate to differences in the degree to which forming the appropriate groupings can be done in a generic manner as opposed to a manner that depends on the specific problem-solving task given to the students. The result is a computing system that can automatically group actions into the hierarchy of Figure 1.

The main contribution of this work is the development and testing of a general methodology for categorizing student problem solving in ELEs. The outcome is a system for classifying student strategies that is automated and so can scale to contexts involving large numbers of students. For classroom use of ELEs, summary information on student strategies and progress may be presented to instructors to allow them to more efficiently provided help and feedback to students. For online uses of ELEs, strategy classification may be used as crucial input to systems that provide hints and feedback or that otherwise customize instruction.

■ RELATED WORK

Our work relates to a large body of literature on inquiry-based learning in educational technological environments. These environments span a wide range of interaction styles, from question-and-answer-based systems that provide a constrained strategy space and a high degree of scaffolding, to Exploratory Learning Environments (ELE) that are open ended and allow students to form hypotheses, run experiments, and evaluate evidence. Here, we discuss some approaches from prior work on analysis of student interactions in computer-based ELEs.

ELEs encourage a "learning by doing" approach that has been shown to increase learning outcomes in students but challenges the analysis of students' learning strategies.[11,12] Recent works have used computational methods to address this challenge. Pedro et al.[13] and Montalvo et al.[14] used machine learning techniques to identify two types of students' planning approaches in microworlds, a simulation-based educational environment.[15] Other works have used data mining techniques to cluster and classify students' learning behaviors in ELEs as either effective or ineffective.[16,17]

The IMMEX project has had considerable success in identifying students' problem solving strategies in a number of domains, including chemistry.[18,19] The problem solving tasks present students with a set of actions, such as performing various chemical tests on an unknown sample. An artificial neural network is then used to classify students into groups that performed similar sets of actions. Further examination of these groups by domain experts is then used to identify the problem solving strategy being employed. The resulting ability to automatically assign strategies to students has enabled studies that examine the effects of collaborative learning and metacognitive interventions on student's strategies.[20−23] The classification algorithm considers only the set of actions carried out by the student and not the order of the actions. This is sufficient for the problem solving tasks being analyzed, since the actions correspond directly to subgoals (e.g., perform a flame test). For the virtual lab ELE considered here, inferring problem-solving strategy necessitates grouping of raw actions (e.g., pouring a solution from one flask to another) into more meaningful sets of actions (e.g., carrying out a titration), which involve parameter (e,g, which solutions are being mixed) and temporal constraints. This leads to the plan recognition approach[24] used here to group *actions* into *activities*.

Given the complexity of analysis in ELEs, the tools mentioned above work post hoc and generate reports and analysis to teachers based on students' complete interaction histories with the software. A notable exception is the work by Noss et al.,[25] who designed a student tracking tool for

DOI: 10.1021/ed500531a
J. Chem. Educ. XXXX, XXX, XXX−XXX

improving algebra generalization in an ELE. The tracking tool visualizes "landmarks"' in real-time which occur when the system detects specific actions or repetitive patterns carried out by the student.

Within the realm of more constrained e-learning systems, the FORMID-Observer monitors students' activities in simulation-based work sessions with the FORMID-Learner. This allows teachers to specify possible student mistakes that are monitored by the system. Automatic analysis of students' activities is performed and a coloring scheme is then used to convey correct and incorrect activities. The ASSISTment system, which provides exercises and tutoring assistance for students, generates reports that summarize students' performance for instructors.[12] The Student Inspector aims to support teachers in distance learning contexts[13] by displaying reports on both individual and group performance. This allows instructors to identify individuals who are performing especially well or who may need more support. Reports on learner's misconceptions, identified through use of the system, are also generated.

Lastly, we consider past work in recognizing students' activities within virtual laboratories for chemistry education in order to provide hints and feedback to students.[26,27] These works rely on recognizing when the simulation has been put into a state that suggests students have met a prespecified subgoal of the problem solution. For example, using the virtual lab to measure the enthalpy of a reaction may be broken into two subgoals: mixing the reactive substances and interpreting the observed temperature change. Once a reactive mixture has been made, the help system assumes the first subgoal has been met and switches hints and feedback to the support of data analysis. The current work goes beyond consideration of only the current state of the simulation by detecting a series of actions that correspond to a problem-solving strategy.

■ METHODOLOGY

Data on Problem Solving

The ChemCollective Virtual Lab (vlab) is an exploratory learning environment with a flexible interface that allows students to design and carry out a wide variety of experiments.[28] In this study, students were given the following task:[24]

You are given four substances A, B, C, and D that are known to react in some weird and mysterious way (an oracle relayed this information to you within a dream). Design and perform virtual lab experiments to determine the reaction between these substances, including their stoichiometric coefficients. The "stockroom" of the virtual lab contains 1.0 M ("molar" concentration unit) solutions of each of these species. A form is provided at the bottom of the page for you to enter their reaction, for example, 2A + B → C + 2D. You are given three chances to submit an answer before having to reload the web page, and so receive a new randomly generated chemical reaction.

This task gives students practice with reaction stoichiometry that complements the practice provided by typical textbook problems. The vlab collects a detailed log of all student actions. The log files considered here were randomly selected from a population of students taking second semester chemistry at an R1 institution during the years 2006 and 2008. Students were assigned the problem as part of a graded homework assignment. There were 156 (82 female, 74 male) students in 2006 and 150 (89 female, 61 male) in 2008, totaling 306 students. Here, we selected 99 log files from those in which the student successfully solved the problem within the three allowed attempts.

Hierarchical Grouping of Actions

The following processes were used to group actions into the hierarchy of Figure 1. As one proceeds up the hierarchy, the degree to which the appropriate groupings rely on the specific problem-solving task increases. The processes used to form the first level of groupings, that of actions into activities, are based on preexisting tools for visualizing students' activities in the ChemCollective Virtual Lab[24] that have been made available to teachers and students using the software.

Actions are the basic forms of interaction, such as pouring 10 mL of solution from one flask into another. At this lowest level, the *temporal visualization* tool (Figure 2) provides a convenient

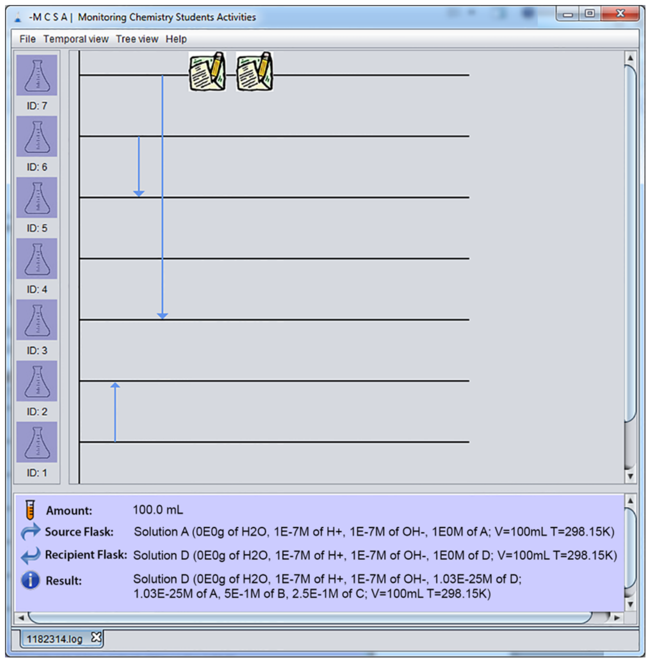

Figure 2. Temporal log file visualization using arrows to indicate source and recipient flasks for three pairwise mixes. Clicking an arrow loads the lower panel with the details for that action.

means for viewing actions in sequence. This tool is generic and will apply across all problems types. Each action taken by the student is drawn over a time-line. The vertical axis displays the flasks used when solving the problem. The horizontal axis shows student actions ordered by time. Each transfer of contents between flasks is shown as an arrow connecting the source flask (i.e., the flask from which solution is being poured) to the recipient flask, with thicker arrows indicating larger transfer volumes. Clicking on an arrow causes details regarding the transfer to be shown in the bottom information panel.

Activities are sets of actions that correspond to common laboratory procedures. A fairly general means of dealing with activities is also feasible because activities relate only to laboratory procedures which cross many problem types. Here, an algorithm is used to group actions into activities, with the resulting groupings being presented via a *tree visualization* tool (Figure 3). The actions are arranged into a tree structure, with the highest level of the tree corresponding to the "activities" of our hierarchical grouping. Because a log file may contain many actions, this tree structure helps significantly with the analysis.

DOI: 10.1021/ed500531a
J. Chem. Educ. XXXX, XXX, XXX−XXX

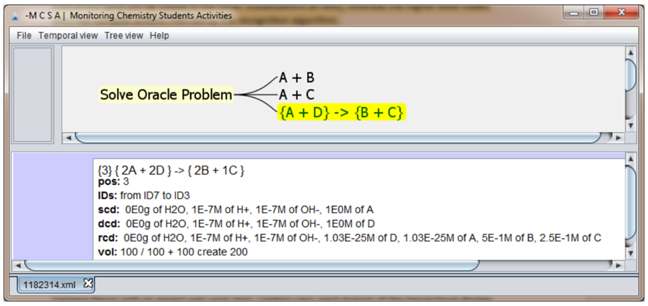

Figure 3. Tree visualization of the log file shown in Figure 2. Branches correspond to "activities" of the hierarchical grouping of Figure 1. Green text color indicates A and D react to produce B and C. Selecting a portion of the tree (yellow highlight) shows additional information in the lower panel.

The goal is for the higher levels of the tree to summarize the actions sufficiently well that it is not typically necessary to drill down to lower levels. The grouping of actions into activities is done using a plan-recognition algorithm that is not problem specific.[24] The algorithm receives as input students' complete interaction sequence with vlab, as well as a grammar describing two types of pouring activities (mixing two substances into the same destination flask, and mixing two substances in a destination flask using a pipet or other intermediate flask). The algorithm constructs a hierarchy by matching the log of the student to activities in the grammar. These activities themselves can then proceed to form new activities in the next iteration, capturing the modular nature by which students solve vlab problems. This approach is successful at grouping actions into activities, even when the actions being grouped are not performed sequentially by the student.

Although the grouping of actions into activities is done using an algorithm that is not problem specific,[24] problem-specific information is displayed as text labels on the nodes, to aid interpretation of students' work. The labels for the vlab problem considered here list the chemical species that were added to the flask followed by the species that were present after the reaction. As compared to the temporal visualization of Figure 2, the tree visualization of Figure 3 makes it easier to group actions into strategies. In Figure 3, it is more apparent that the student made three pairwise mixes and stopped when the third mix led to a reaction. As discussed below, this corresponds to an "expert pairwise mix" strategy.

Strategies are sets of activities carried out with a specific intent regarding problem solving. The grouping of activities into problem-solving strategies is sufficiently problem specific that it is difficult to imagine a general algorithm for identifying strategies. Instead, our process uses human analysis of log files to gather the information needed for a problem-specific algorithm. Domain experts use the visualization tools discussed above, for actions and activities, to develop a coding scheme for problem solving strategies. This coding scheme is then translated into a rule-based computing system that can automatically detect the presence of a strategy in a log file. The result is an automated system for grouping actions into strategies.

Solution paths are sets of strategies that, taken together, allow one to solve a given problem. The solution path is taken as simply the set of all strategies identified in a particular log file.

Evaluation of the Resulting Automated System

To evaluate the resulting computing system, the log files were separated into a training set of 30 files and a test set of 69 files. The training set was used to develop the coding scheme for strategies and create the rules for the automated system, as discussed in the Hierarchical Grouping of Actions section above. The system was then evaluated by comparing human generated with computer generated strategy codes on the test set.

RESULTS

Coding Scheme and Rules for Strategies

Examination of the 30 logs files in the training set led to the strategies summarized in Figure 4. A typical solution path

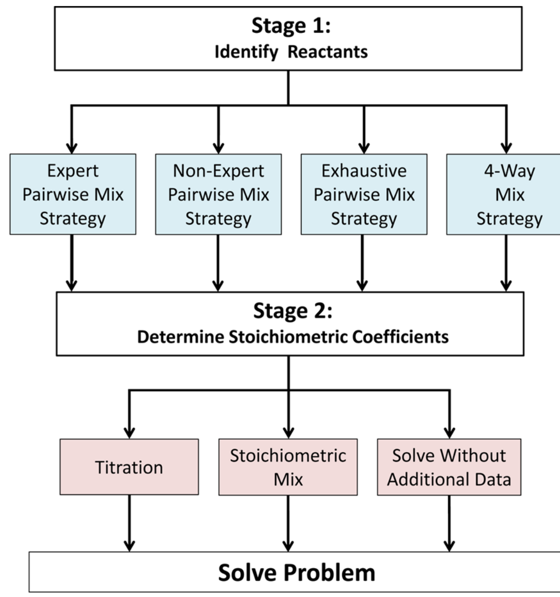

Figure 4. Coding scheme (blue and red boxes) for attaching strategies (problem solving intent) to sets of activities. A solution path typically combines strategies from the two stages shown here.

exhibits the two stages shown in Figure 4: discovering which species react with one another (top) and determining the full stoichiometry of the reaction (bottom).

For the first stage (top of Figure 4), the most common strategy is to mix pairs of species (A with B, A with C, etc.). In some cases, students stop mixing pairs when a reaction occurs because this is sufficient to identify the reactants. We label this as an "expert pairwise" strategy because the student is attending to the results as they proceed. In the "exhaustive pairwise" strategy, all six possible pairwise mixes are created, even if a reaction occurs before the last pairwise mix. (For cases where a reaction is not observed until the student makes the last possible pair, it is not possible to distinguish expert from exhaustive, and we label it as exhaustive.) "Non-expert pairwise mix" refers to pairwise strategies that do not fit into either the expert or exhaustive strategies, or in which other actions are interspersed. An additional strategy seen in the first stage of problem solving is a "four-way mix" in which all four species are mixed in a single flask.

In the second stage (bottom of Figure 4), the solution process shifts to determining the products and the stoichiometric coefficients. Any mixture in which a reaction occurred

DOI: 10.1021/ed500531a
J. Chem. Educ. XXXX, XXX, XXX−XXX

has sufficient information to determine the stoichiometry. However, students often perform additional experiments, which we will illustrate for the reaction A + 2B → C + 2D. In the "titration" strategy, one of the reactants is added incrementally to the other reactant. This is a convenient way to determine the ratio of the stoichiometric coefficients for the reactant species, e.g. if it takes 200 mL of 1 M B to totally consume the A present in 100 mL of 1 M A, then the reaction stoichiometry between B and A must be 2 to 1. In the "stoichiometric mix" strategy, the reactants are mixed in stoichiometric proportions, for example, 200 mL of 1 M B are mixed with 100 mL of 1 M A. A student may be carrying out a stoichiometric mix in order to test a hypothesis regarding the reactant stoichiometry, for example, to confirm that B and A react in 2:1 proportions. Another possible motivation for a stoichiometric mix is to aid in the identification of product stoichiometry because the reactants are totally consumed leaving only the products.

The next step is to convert the above human coding scheme into rules that can be implemented in an automated computing system for detecting the presence of the strategies of Figure 4 in a log file. The rules used in this study are listed in the "Supporting Information".

Because the log files contain only the student actions in the virtual lab and submissions of answers through the web form, the domain experts must infer the process used by the student to interpret the experimental data. Many log files contain actions that the domain experts do not attribute to a principled approach to solving the problem. The presence of such actions is evidence of students engaging in the type of exploratory behavior that ELEs are meant to encourage. When a set of activities in a log file satisfies the criteria for a particular strategy, that log file is labeled as containing that strategy. We note that the detection of a strategy in a log file does not guarantee the student was intentionally engaging in that strategy as opposed to simply engaging in exploration.

The chemical concepts that are targeted by this virtual lab activity include those of reaction stoichiometry and limiting reagents. To illustrate potential connections between these concepts and the above strategies, consider the following example, from our classroom observations, of how some students interpret data from a four-way mix of equal volumes of all species (Table 1). Factoring in only the effects of dilution

Table 1. Tabular Format That Classroom Observations Found Some Students Use to Interpret Data from a Four-Way Mix Strategy

	A	B	C	D
Initial	0.25 M	0.25 M	0.25 M	0.25 M
Change	−0.125 M	+0.125 M	−0.125 M	+0.125 M
Final	0.125 M	0.375 M	0 M	0.375 M

leads to *initial* concentrations of 0.25 M. Deviations of the *final* concentrations (observed in the lab) from 0.25 M must be due to *changes* due to the chemical reaction. The sign of the changes indicate that A and C react to produce B and D. The ratios of the changes give the stoichiometry of the reaction as A + 2C → B + D. Although we can detect the presence of a four-way mix in a log file, we cannot infer that they are using this approach to analyze the data.

Distribution of Identified Strategies

The frequency of pairwise strategies is shown in the top row of Table 2. There are 17 logs that do not contain a pairwise strategy and instead contain only a four-way mix. Of the pairwise strategies, expert and nonexpert have nearly equal frequencies, 31 and 30 logs respectively, with exhaustive being less common, occurring in 20 logs. The last column shows that roughly half of the logs contain a four-way mix and that roughly half of the logs contain one or more of the Stage 2 strategies. The lower portion of Table 2 examines how the frequency of the Stage 2 strategies interacts with the strategy used in Stage 1. For students using one of the pairwise strategies in Stage 1, a few trends emerge in Stage 2. Students using a nonexpert approach to the pairwise strategy were more likely than average to carry out a four-way mix (73%) and to carry out at least one Stage 2 strategy (67%). This suggests a highly explorative approach that gathered substantial experimental information before solving the problem. Students using an expert approach to the pairwise strategy were unlikely to perform a four-way mix (6%) and used a Stage 2 strategy (48%) with a frequency comparable to that of the entire population (59%). This suggests a principled and efficient approach to solving the problem. These interactions between Stage 1 and Stage 2 strategies suggests the combination of all strategies present in a log file, that is, the "solution path" at the top of our hierarchical grouping, may provide useful insight into student problem solving.

Table 2. Distribution of Strategies Detected in 98 Log Files

	Exhaustive Pair	Expert Pair	Non-Expert Pair	No Pairwise	Total
Stage 1 Strategies					
Pairwise Strategy[a]	20	31	30	17	98
Four-Way Mix[b]	11 (55%)	2 (6%)	22 (73%)	17 (100%)	52 (53%)
Stage 2 Strategies					
Titration[b]	3 (15%)	9 (30%)	11 (37%)	4 (24%)	27 (28%)
Stoichiometric Mix[b]	4 (20%)	7 (23%)	13 (43%)	3 (18%)	27 (28%)
No Phase 2 Strategy	13 (65%)	16 (52%)	10 (33%)	11 (65%)	50 (51%)

[a]Number of log files in which each pairwise strategy was detected. [b]Co-occurrence of additional strategies with the various pairwise mix strategies. For instance, of the 30 logs that contained a nonexpert pair strategy, 22 logs (73%) also contained a 4-way mix and 11 logs (37%) also contained a titration. (Note that some logs contain multiple nonpairwise strategies.)

Validating the Rule Based System

The automated rule-based system for classifying student strategies was evaluated through the following empirical study. We sampled 99 student logs of varying sizes and complexities. Thirty of these logs formed the training set, which were visualized, using the tools of Figures 2 and 3, to develop the strategies of Figure 4. All logs were then hand-coded for the presence of the strategies. Two coders worked together, with any disagreements being negotiated to an agreed upon final code. This process was based on a holistic examination of the entire problem-solving sequence and the coders' best attempt to infer student intent. The distribution of students' strategies is shown in Table 2.

The rule based system was then applied to each log file (See Supporting Information). Since codes were based on a visual presentation of data in the log file, it is feasible that the principles being applied by the human are well summarized by

DOI: 10.1021/ed500531a
J. Chem. Educ. XXXX, XXX, XXX−XXX

the rules, and so are built into the "expert system". To evaluate this, we compared the strategies that were outputted by the program to those determined by the coders.

We present the results separately for recognizing strategies for Stage I of the problem and for Stage II, as described in Figure 4. For Stage I, there was agreement for 91 of the 98 log files (93% agreement). Examinations of the logs revealed a variety of reasons for the disagreements. In some cases, nonproductive actions (three-way mixes, or repeated mixing of pairs) were interspersed with an expert or exhaustive pairwise mix in a manner that caused the rules to misidentify the type (expert, exhaustive, or nonexpert) of pairwise mix strategy. In other cases, pairwise mixes following a four-way mix were assigned to a stage II strategy while the human coder labeled them as a stage I pairwise strategy.

For stage II, there was agreement for 87 of the 98 files (89%). In all cases, disagreements arose from the rule-based system assigning a phase II strategy to actions that the human coders had labeled as part of phase I. For example, a log was labeled by humans as containing only a four-way mix followed by an exhaustive pair, whereas the rule-based system agreed on these phase I strategies but also labeled part of the exhaustive pair mixes as a stoichiometric mix. Although it may be possible to extend the rule set to address some of these cases, this was viewed as likely to lead to overspecialization on the particular set of logs being used in the empirical evaluation. An accuracy of 89% for these logs suggests the process developed here is capable of creating a system that can usefully automate the classification of problem solving strategies within students' solution paths.

■ DISCUSSION

This paper presents a staged process for automatically classifying student problem solving strategy in an open-ended virtual laboratory. The process used here is based on grouping actions in a hierarchical fashion, moving from generic actions and activities toward problem-specific strategies and solution paths. This approach has the potential to apply across a wide variety of problem types and learning environments.[29] An automated approach to analyzing student problem-solving in exploratory learning environments has a number of important potential uses. Instructors can use this information in classroom settings to better monitor progress of a large group of students and provide help as needed. Computer tutoring systems can also take advantage of this information to provide interventions that are based on the students' current problem solving strategy. The form of scaffolding that maximizes learning remains to be explored; however, a key enabling technology for any such approach is the ability to infer problem-solving intent from student actions.

The process developed here has a number of upfront costs and limitations. The visualization tools of the "Hierarchical grouping of actions" section have an upfront cost associated with learning how to interpret the graphs. Past work indicated that the tree visualization (Figure 3) is preferred over the temporal visualization (Figure 2) by only those 50% of instructors who, in a 1 h session, understood how to interpret the tree.[24] Because the strategy detectors are problem-specific, there is also an upfront cost with developing the rules used to label strategies. Development of these rules is the primary bottleneck in extending this approach to other problem types. For the problem considered here, examination of an individual log file took an average of about 10 min each. Examination of 30 files was sufficient to develop the coding scheme of Figure 4, but considerable time was also spent reflecting on what was learned from the files and distilling this into a coding scheme. Creating rules that capture the intent of the coding scheme also took considerable time. Our current best estimate of the time required to extend this approach to a new student task is a few weeks for a team of two domain experts. The accuracy of the strategy classification is limited by the use of only actions occurring in the virtual lab or the web forms used to report the answers, as opposed to what students are thinking or writing down outside the computer environment. Nevertheless, once developed, this system provides a means to automate the detection of problem solving strategy, which yielded 89% accuracy for the example considered here when compared with human assigned codes.

Looking forward, the process developed here has the potential to lead to scalable instruments for formative assessment of online inquiry activities. Such activities may be viewed as coupling conceptual understanding of domain content with science practices such as experimental design and data interpretation.[2,3] Concept inventories provide a scalable means for assessing conceptual understanding.[30] The current work may provide a scalable means for assessing science practices. The process developed here has some parallels to the process used to develop concept inventories. Both processes begin with a free form log. Here, this is a raw list of actions in the virtual lab, whereas for concept inventories, it is interview transcripts. In both cases, the goal is to achieve scalability by automating the process of categorizing student reasoning. For concept inventories, the categories are reflected in the choices made available to the student in a fixed response format. Here, the categories are reflected in the strategy codes of Table 2. The automated means to assign strategy codes developed here, therefore, may provide an approach to formative assessment of science practices that is scalable and can be applied consistently in diverse contexts.

■ ASSOCIATED CONTENT

Supporting Information

A detailed list of the rules used to automate the process of detecting problem solving strategies in student interaction log files from the virtual lab. This material is available via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: yaron@cmu.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The research reported here was supported in part by the Pittsburgh Science of Learning Center (National Science Foundation grant number SBE-0354420), the National Science Foundation (DUE-1123355), and the Israeli Science Foundation grant no. 1276/12.

■ REFERENCES

(1) Chamely-Wiik, D. M.; Haky, J. E.; Louda, D. W.; Romance, N. SQER 3: An Instructional Framework for Using Scientific Inquiry To Design Classroom Demonstrations. *J. Chem. Educ.* **2014**, *91* (3), 329–335.

DOI: 10.1021/ed500531a
J. Chem. Educ. XXXX, XXX, XXX–XXX

(2) Talanquer, V. Chemistry Education: Ten Facets To Shape Us. *J. Chem. Educ.* **2013**, *90* (7), 832−838.
(3) Cooper, M. M. Chemistry and the Next Generation Science Standards. *J. Chem. Educ.* **2013**, *90* (6), 679−680.
(4) Amershi, S.; Conati, C. Automatic Recognition of Learner Groups in Exploratory Learning Environments. In *Intelligent Tutoring Systems SE - 46*; Ikeda, M., Ashley, K., Chan, T.-W., Eds.; Lecture Notes in Computer Science; Springer: Berlin Heidelberg, 2006; Vol. *4053*, pp 463−472.
(5) Cocea, M.; Gutierrez-Santos, S.; Magoulas, G. D. Adaptive modelling of users? strategies in exploratory learning using case-based reasoning. *Knowledge-Based and Intelligent Information and Engineering Systems* **2010**, *2*, 124−134.
(6) Moore, E. B.; Herzog, T. A.; Perkins, K. K. Interactive simulations as implicit support for guided-inquiry. *Chem. Educ. Res. Pract.* **2013**, *14* (3), 257−268.
(7) Vanlehn, K.; Lynch, C.; Schulze, K.; Shapiro, J. A.; Shelby, R.; Taylor, L.; Treacy, D.; Weinstein, A.; Wintersgill, M. The Andes Physics Tutoring System: Lessons Learned. *Int. J. Artif. Intell. Ed.* **2005**, *15* (3), 147−204.
(8) diSessa, A. A. Toward an Epistemology of Physics. *Cogn. Instr.* **1993**, *10* (2), 105−225.
(9) Bodner, G. M.; McMillen, T. L. B. Cognitive restructuring as an early stage in problem solving. *J. Res. Sci. Teach.* **1986**, *23* (8), 727−737.
(10) Schoenfeld, A. H. *How We Think: A Theory of Goal-Oriented Decision Making and its Educational Applications*; Studies in Mathematical Thinking and Learning Series; Taylor & Francis: Oxford, U.K., 2010.
(11) Ben-Naim, D.; Marcus, N.; Bain, M. Visualization and analysis of student interactions in an adaptive exploratory learning environment. In *1st International Workshop on Intelligent Support for Exploratory Environments*; CEUR, Aachen University: Aachen, Germany, 2008.
(12) Romero, C.; Ventura, S. Educational Data Mining: A Review of the State of the Art. *IEEE Trans. Syst. Man, Cybern., Part C: Appl. Rev.* **2010**, *40(6)*.
(13) Pedro, M. A. S.; Baker, R. S. J.; Montalvo, O.; Nakama, A.; Gobert, J. D. Using Text Replay Tagging to Produce Detectors of Systematic Experimentation Behavior Patterns. *Proceedings of the 3rd International Conference on Educational Data Mining*; International Educational Data Mining Society: Massachusetts, 2010; pp 181−190.
(14) Montalvo, O.; Baker, R. S. J. D.; Pedro, M. A. S.; Nakama, A.; Gobert, J. D. Identifying Students ' Inquiry Planning Using Machine Learning. *3rd International Conference on Educational Data Mining*; International Educational Data Mining Society: Massachusetts, 2010; pp 141−150.
(15) Edwards, L. D. Embodying mathematics and science: Microworlds as representations. *J. Math. Behavior* **1998**, *17* (1), 53−78.
(16) Amershi, S.; Conati, C. Combining unsupervised and supervised classification to build user models for exploratory learning environments. *J. Educ. Data Min.* **2009**, *1* (1), 1−54.
(17) Kardan, S.; Conati, C. A framework for capturing distinguishing user interaction behaviours in novel interfaces. *EDM 2011 - Proceedings of the 4th International Conference on Educational Data Mining*; International Educational Data Mining Society: Massachusetts, 2011; pp 159−168.
(18) Cooper, M. M.; Sandi−Urena, S.; Stevens, R. Reliable multi method assessment of metacognition use in chemistry problem solving. *Chem. Educ. Res. Pract.* **2008**, *9* (1), 18.
(19) Sandi-Urena, S.; Cooper, M. M.; Gatlin, T. A.; Bhattacharyya, G. Students' experience in a general chemistry cooperative problem based laboratory. *Chem. Educ. Res. Pract.* **2011**, *12* (4), 434.
(20) Cox, C. T., Jr.; Cooper, M. M.; Pease, R.; Buchanan, K.; Hernandez-Cruz, L.; Stevens, R.; Picione, J.; Holme, T. Advancements in curriculum and assessment by the use of IMMEX technology in the organic laboratory. *Chem. Educ. Res. Pr.* **2008**, *9* (2), 163−168.
(21) Cooper, M. M.; Cox, C. T.; Nammouz, M.; Case, E.; Stevens, R. An Assessment of the Effect of Collaborative Groups on Students' Problem-Solving Strategies and Abilities. *J. Chem. Educ.* **2008**, *85* (6), 866.
(22) Beal, C.; Stevens, R. Improving Students' Problem Solving in a Web-based Chemistry Simulation Through Embedded Metacognitive Messages. *Technol. Instr. Cogn. Learn.* **2011**, *8* (3), 255−271.
(23) Sandi-Urena, S.; Cooper, M.; Stevens, R. Effect of Cooperative Problem-Based Lab Instruction on Metacognition and Problem-Solving Skills. *J. Chem. Educ.* **2012**, *89* (6), 700−706.
(24) Amir, O.; Gal, Y. Plan Recognition and Visualization in Exploratory Learning Environments. *ACM Trans. Interact. Intell. Syst.* **2013**, *3*, 16:1−16:23.
(25) Noss, R.; Poulovassilis, A.; Geraniou, E.; Gutierrez-Santos, S.; Hoyles, C.; Kahn, K.; Magoulas, G. D.; Mavrikis, M. The design of a system to support exploratory learning of algebraic generalisation. *Comput. Educ.* **2012**, *59* (1), 63−81.
(26) Borek, A.; McLaren, B.; Karabinos, M.; Yaron, D. How Much Assistance Is Helpful to Students in Discovery Learning? In *Proceedings of the fourth European conference on technology enhanced learning, learning in the synergy of multiple disciplines (EC-TEL 2009)*; Cress, U., Dimitrova, V., Specht, M., Eds.; Springer-Verlag: Berlin/Heidelberg, 2009.
(27) Davenport, J. L.; Rafferty, A. N.; Timms, M. J.; Yaron, D.; Karabinos, M. ChemVLab+: Evaluating a Virtual Lab Tutor for High School Chemistry. In *The Proceedings of the 2012 International Conference of the Learning Sciences*; International Society of the Learning Sciences: Sydney, 2012.
(28) Yaron, D.; Karabinos, M.; Lange, D.; Greeno, J. G.; Leinhardt, G. The ChemCollective—virtual labs for introductory chemistry courses. *Science* **2010**, *328* (5978), 584−585.
(29) Gal, Y.; Reddy, S.; Shieber, S. M.; Rubin, A.; Grosz, B. J. Plan recognition in exploratory domains. *Artif. Intell.* **2012**, *176* (1), 2270−2290.
(30) Luxford, C. J.; Bretz, S. L. Development of the Bonding Representations Inventory To Identify Student Misconceptions about Covalent and Ionic Bonding Representations. *J. Chem. Educ.* **2014**, *91* (3), 312−320.

DOI: 10.1021/ed500531a
J. Chem. Educ. XXXX, XXX, XXX−XXX