Assessment of Group and Individual Learning through Intelligent Visualization Workshop (AGILeViz)

Supplementary Proceedings of the 13th International Conference of Artificial Intelligence in Education.

Marina del Rey, CA. USA.

July 2007
Assessment of Group and Individual Learning through Intelligent Visualization Workshop (AGILeViz)

AIED 2007

Eric Hamilton and Andrew Hurford

Workshop Committee:


Purpose of the Workshop: This workshop will present, discuss, and explore new ways of providing educators with sophisticated, computer-generated visualizations of learning by groups and individuals in order to improve teaching and learning. The organizers of the workshop are motivated by the convergence of several factors that bear directly on the AIED vision to understand and encourage novel, technology-rich learning and learning tools:

- Educational innovators are using technology to understand and propel learning in ways that dwarf the limited paradigms of learning and assessment that dominate educational practice.

- Current assessment tools not only are highly constrained in the constructs that they measure, but the representational systems used to communicate student progress on these assessments are also very limited. Representational systems used are inadequate for supporting thinking about the kinds of learning and problem-solving behaviors that are increasingly important to assure.

- Interactive visualization techniques for large quantities of heterogeneous data have now reached a state of maturity that make them ripe for applications in data-driven analysis and decision making for education.

The workshop is based on a conjecture that complex visualization systems may be applied to analyses of learning and in support of learning to create a new class of tools and environments for promoting collaborative learning and learning research. Our goal is to nurture the nascent area of applying complex graphical and intelligent systems for analysis and assessment of complex individual and group learning. This half-day workshop will feature presentations by the committee on the state of the art in their respective research and will respond to questions from the audience. After four presentations outlining the scope of possibilities in innovative assessment and representation schemes, the participants and committee will form small groups brainstorming, collecting, and recording ideas, suggestions, and questions.
Table of Contents

A Visual Concept Mapping Medium to Open Student and Group Models  
Urko Rueda, Ana Arruarte, Jon Elorriaga……………………………………………………p. 4

Assessment of Group and Individual Learning Through Visualization (AGILEeViz)  
Eric Hamilton……………………………………………………………………………………p. 8

Visualizing Student Learning and Understanding in Introductory Physics  
Troy Hacker……………………………………………………………………………………p. 18

Making Sense of Complex Learner Data  
Christopher Brooks, Wengang Liu, Collene Hansen, Gordon McCalla, Jim Greer…p. 28

The Social Affordance of a Paper  
Tiffany Y. Tang and Gordon McCalla…………………………………………………………p. 34
A Visual Concept Mapping Mapping Medium to
Open Student and Group Models

Rueda, U., Arruarte, A., Elorriaga, J. A.
University of the Basque Country

Abstract.
Inside the computer based education community there is a need to open Student
Model to the users (e.g. teachers and students) in order to promote reflection, in
students, and allow inspection to the teachers. In this paper we present DynMap+, a
visual medium that allows the inspection of Individual and Group Student Models
coming from different sources. Several visual mechanisms are used to open those
models adequately. DynMap+ is valid for visualising not only the last state of the
models but also the evolution through the learning sessions.

Introduction

One of the most important issues in the educational agenda is to achieve students’ learning
objectives effectively. That means to help the student acquire the required level of
knowledge and skills in the subject domain. Thereby, it is necessary to adapt the teaching to
each student particular needs. It is commonly agreed that, for adaptation, some kind of
student representation is needed [1], [2]. In the Artificial Intelligence in Education area, the
mechanism to represent the state and evolution of student learning are Student Models.
Several ways of representing Student Models can be found in the literature. However, all of
them share a main idea: to collect the information related to the student that influences in
his/her learning such as the level of knowledge, the acquired skills, the learning objectives,
the learning preferences, etc. Once a Student Model is available, the teaching/learning
system uses it for adaptation.

Student Models have been criticised for not facilitating the access to the data they
contain. To solve this lack, the AI-ED community has made efforts to open Student Models.
An Open Student Model is a student representation designed for allowing inspection. In
addition, for a better understanding, graphical representations are appropriate. Cook and Kay
[3] make one of the first attempts for graphically representing the Student Model; they use a
mixture of text and conceptual trees based diagrams. Morales et al. [4] present a table-like
graphical representation format of the Student Model for a sensory-motor task in which
every row represents a rule. LeMoRe1 (Learner Modelling for Reflection) groups several
examples of Open Student Models.

The Open Student Models can be classified into three main groups: raw data models,
visual models and decision support models. A raw data model can be defined as a direct
opening of the internal representation. The problems of those models are mainly due to the
difficulty in their understanding. This does not happen in a visual model. A visual model
converts the internal representation to a graphical conceptualization that allows a better
recognition of the learning aspects. Moreover, in a decision support model the user can
obtain greater advantages. A decision support model can be defined as a visual
representation medium that allows the user to make pedagogical decisions based on the

1 http://www.eee.bham.ac.uk/bull/lemore/examples.html
learning characteristics displayed. To make Open Student Models more useful for the educational community, they should be able to be opened following the third main group of models: Decision Support Models. In order to achieve it, some design principles should be defined.

*Bull and Nghiem* [5] claims that the use of simple Student Models, easy to represent in different ways, e.g. graphical and tabular, allows teachers and students to better understand students’ learning of a target domain. Nevertheless, before they could make decisions from the displayed information, a previous knowledge of the representation must be achieved. In order to bypass this difficulty, a proper visualization medium is needed.

A Concept Map is a useful resource for opening Student Models. Since Novak [6] placed Concept Maps in the educational agenda, they have been broadly used as a medium to represent and interchange knowledge [7]. A Concept Map is a graphical way of representing and organizing knowledge. It is comprised of nodes and links, arranged in some order to reflect the domain being represented. Nodes symbolize concepts, and links represent relationship between concepts; both concepts and links are labelled and may be categorized. The use of Concept Maps can help to improve the understanding of the information contained in Student Models.

Next sections focus on the following aspects. First, some design principles to open Student Models through Concept Maps are introduced. Next, DynMap+ is presented. It allows visualizing Decision Support Models and even allows inspecting Individual and Group Student Models. Finally, some conclusions are pointed out.

### 2. How to open Student Models through Concept Maps: Design principles

Some considerations can be pointed out when opening Student Models as Concept Maps. A usual failure spot comes from the scalability problem of Concept Maps. When the number of nodes increases, the readability of the Concept Map decreases and the understanding ability comes to a failure. To avoid the scalability problem, the visual medium must offer enough visual mechanisms to meet Dimitrovas’ criteria about compressibility, effective inspection and cognitive overload reduction [8].

Also, the original Novak’s definition of Concept Maps should be extended. In the actual era of technology, Concept Maps can benefit from applying graphical resources over them. For example, images, visual symbols, and variations in the drawing of the Concept Map graphic elements allow acquiring a better readability of the information displayed.

### 3. DynMap+: A medium to visualize Decision Support Models

DynMap+ is a visual mechanism based on Concept Maps that allows opening Student Models coming from different sources [9]. Sources vary from student data collected in a classical teaching context to a teaching/learning system with its own Student Model representation. In order to capture external data, it must be translated to the internal representation of DynMap+. DynMap+ is able to open the learning characteristics related to both individual students and groups. The knowledge of students is represented following an overlay approach [10]. Thus, the knowledge the student has is viewed as a subset of the whole learning domain. As the knowledge of students’ changes over the learning sessions, the Student Models should support this dynamic behaviour. DynMap is able not only to maintain the last state of knowledge of students, but also the knowledge evolution during the learning sessions.

Figure 1 shows a screenshot of DynMap+ in the context of a Computer Security
course at the University of the Basque Country. Two views of the same student group are displayed. In the upper side, the user can inspect the learning evolution of the group based on a graphical representation of the domain that includes structural and pedagogical relationships. In the lower side the skills worked by an individual student of the group are displayed.

In order to increase the representation ability of the Concept Maps, several graphical resources are used. The node form represents the type of the content (e.g. modules of the domain, units in each module, learning activities, etc.). Outline thickness of the border represents the student’s level of knowledge about a concept. In order to determine the thickness the number of learning activities and their marks are taken into account. The thicker the border is, the more knowledge the student has about the concept. A dashed border represents that the concept is not completely achieved, e.g. the performance of the student when working with the learning activities related to the concept has gone beyond a threshold, and otherwise the line will be continuous. Flags are used to show when the
concept has attached learning activities (for example, lectures, labs, assignments). Thus, the Concept Map shows the type of learning activities used in each part of the domain. Shadowed nodes show whether the node are contracted (shadowed) or not. A contracted node contains a subset of the Concept Map related to the node. Finally, a green rectangle is used to highlight new nodes in the Student Model (student’s knowledge changes). Also, for Group Student Models attenuation in the colour of the nodes is used to visually display the contribution of students learning over the concepts. The more attenuated the fewer students that have inverted efforts in the learning of the concept. Those graphical resources combined with other visualization mechanisms like views and filters solve the potential scalability problems of this approach.

4. Conclusions

In this paper, DynMap+, a tool for supporting Group Student Models has been presented. The tool represents not only the last state of a model but also its evolution during the learning sessions. The models that follow the overlay approach are graphically visualized by means of Concept Maps. With this tool, useful information is opened for inspection, allowing teachers to make teaching decisions (e.g. new learning activities) in order to improve the students learning. Therefore, the tool can be classified inside the Decision Support Systems as it externalizes the state of knowledge of individuals and groups of students for decision making purposes. A direct application of the tool is the formation of group of students. By suggesting appropriate group arrangements.

References

Assessment of Group and Individual Learning Through Visualization (Agileviz) In Education

Eric R. Hamilton
Director, Center for Research on Teaching and Learning, US Air Force Academy, USA

Abstract: This paper makes two arguments en route to the proposition that it may be possible, in a relatively short time frame, to effect a quantum jump in the utility of existing learner assessments via the use of integrative computer graphics. Sophisticated visualization systems may be applied to learner assessments to create a new class of instruments that more effectively highlight complex phenomena in learning. The first argument is that despite their sophistication, modern assessment tools provide limited views into learner cognition and achievement. The second is that future learning environments will entail important educational dynamics that are not accommodated by current assessment approaches. This leads to a conjecture that complex visualization systems may be applied to learner assessments to create a new class of instruments can offer knowledge and insights into complex phenomena in learning.

Keywords: assessment, complex reasoning, graphics, assessment visualization, eportfolios

Introduction

The purpose of this paper is to stimulate interest and conversation in an area of significant potential for assessment systems in future education environments. Currently there is a serious disconnect between the affordances of assessment practices, how assessment results are represented, and the kinds of information that should be made available to and usable by teachers and professors. This paper provides two arguments that lead to the following proposition: sophisticated visualization systems may be applied to learner assessments to create a new class of instruments that more effectively highlight complex phenomena in learning. The first argument is that despite their sophistication, modern assessment tools provide limited views into learner cognition and achievement. Formative assessments and those that entail portfolios broaden the range of what can be measured effectively but have inherent limitations related to the time required to organize and to review them. The second argument is that learning environments themselves are undergoing change, and the learning environment of the future will entail important educational dynamics that are not detected by current assessment approaches, further intensifying the need for new and better assessments. These arguments lead to a discussion of the prospects for the use of complex graphics to integrate multiple assessment data sets into single intuitive displays that depict complex learning.
1. Modern assessments provide limited views into learner cognition and achievement

The meme of improving assessment systems, both formative and summative, is a baseline element in dialogs about scholarship and reform in numerous areas [3-7]. The collective sense that typical summative and formative assessments are inadequate is multifaceted. The advent of the Force Concept Inventory (FCI) [8, 9] in the early 1990s and inventories that developed later [10, 11], disclosing significant conceptual difficulties held by students despite apparent success in their coursework, has been one watershed in recognizing and addressing shortcomings in traditional assessment approaches. Another element is the growing awareness that one of the overarching goals of education, to promote the development of complex and multifaceted thinking skills, does not necessarily follow from the summing of individual courses. Recent literature highlighting inadequacies in undergraduate education [12], for example, are consistent with observations that, at least by the Reflective Judgement [13] and other tools, students seem not to develop complex reasoning skills in college. Students do acquire knowledge, but in terms of becoming better thinkers, more capable of integrating multiple frameworks and making complex judgements, the available evidence indicates that students make very little progress [14].

A third element is recognition of the limitations of traditional *summative* testing. Psychometrics furnish tools that give inferential power but in relatively univariate ways. Standardized tests, by way of illustration, are one measure of knowledge in a content area. The statistics from which an instructor can draw reliable inferences about students derive their validity from the mathematics of one variable of analysis. Many of the deficiencies in assessment systems in education are shared in education more broadly and are the subject of the US National Research Council assessment study *Knowing What Students Know* [15].

More sophisticated approaches to educational measurement and extending beyond but incorporating student assessment, such as hierarchical linear modeling [16, 17], can detect variance at nested levels of institution, class and individual, are growing in use, but have generally not been implemented at the undergraduate level. They provide mathematically valid ways not of combining multiple variables for analysis, but of isolating the effects of single variables embedded in larger systems. They do not yield insights on the connected competencies held by students or by groups of students, competencies about which it will be increasingly important for professors to make judgements.

Finally, traditional locally-developed examinations measure against a sampling of knowledge students are expected to acquire in a course. However they often provide no more than an indirect measure of whether a student’s memory decay was completed before or after an examination, or whether cognitive rehearsal behavior prior to the examination (studying) mitigated memory decay longer for one student than for another. Often the midterm or final is just a marker in a shadow race: do the memory traces of course material fade before (lower grade) or after (higher grade) the test, and can studying delay the fade? That is not what education should be about.

Of course, good educators do create instruments that are integrative and go beyond recitation of the kinds of declarative and procedural knowledge that are likely to disappear when not in use. But even well-developed tests generally produce a single valued result, from which instructors are expected to make critical inferences about the enormously complex entity of a student’s development towards professional competence.
2. Learning Environments Of The Future Will Entail Important Educational Dynamics That Are Not Accommodated By Current Assessment Approaches.

Table 1 depicts six principles discussed and published elsewhere [1, 2] that reflect the author’s view of changes that are and will be evident in the emergence of future learning environments. In part, these changes arise from the development of new technologies, but also in relation to demands. The first principle, increasing sightlines in learning environments, reflects the proposition toward which this paper is leading: new visualization tools can enable much greater insights into learning progress. The second principle, increased emphasis on modeling, is responsive to some of the pressures associated with the scarcity of time in curricula for absorbing more content. It suggests that focus on large principles may leverage both more sophisticated reasoning skills and provide different standards for determining what information gets added to an already overflowing curriculum. Further, it provides another motivating illustration for looking at assessment differently.

Table 1. Six Principles of Future Learning Environments [1, 2]

<table>
<thead>
<tr>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased sightlines in the classroom</td>
</tr>
<tr>
<td>Increased emphasis on models and modeling</td>
</tr>
<tr>
<td>Increased connectedness</td>
</tr>
<tr>
<td>Increased “one-to-one-ness” in the classroom</td>
</tr>
<tr>
<td>Increased fluidity of learning context</td>
</tr>
<tr>
<td>Increased interactional bandwidth</td>
</tr>
</tbody>
</table>

focused on how students *mathematize* elements of a problem situation, i.e., impose logical or mathematical understandings or interpretations, and how they make judgments about which parts of a situation are useful to consider with mathematical representations [25]. (In mathematics education research, mathematizing a problem is sometimes used synonymously with the term “modeling.”)

3. The Use of Visualization to Depict Learner Progress

One element or principle of modeling as an approach to teaching is that modeling tends to emphasize just-in-time and need-based learning. The availability of search and
retrieval tools provides an important avenue for acquiring information as it is needed. Modeling also stresses the use of “big ideas” alongside retrievable procedural and technical competencies. This suggests a view of cognition that veers away from information processing models that are binary in character – the view, for example, that an individual either does or does not possess a particular procedural, declarative or strategic competency. Instead, a modeling view emphasizes that ideas and competencies related to using them develop in context, are often patchy or uneven, and become more stable and usable as they take form in models for solving meaningful problems. Such a view begs the question: if ideas and competencies are not acquired on a binary have it/don’t have it basis, if they develop unevenly and in relation to each other and in the context of needing to solve real problems, how should one assess learner progress? Any progress under that view is a complex achievement that simply will not reduce to a single number. Learner progress is a fundamentally relational and context-connected achievement that might be more accurately seen as a dynamic landscape rather than as a single number.

This is exactly the approach developed by Richard Lesh, a prominent researcher in modeling in K12 applied mathematical problem solving. He formulated the construct of “Learning Progress Maps” when a senior scientist at ETS [35]. Learning Progress Maps, or LPMs, can be defined as graphic, n-dimensional, interactive displays that represent student achievements and simultaneously trace their evolution. The LPMs shown in Figure 1 are computer-generated interpretations of early examples of what a group of teachers produced to describe learner progress. Lesh discussed one LPM design activity [36] as follows: “During the process of designing useful LPMs, one distinguishing characteristic of the most effective teachers was the clarity with which they recognized the small number of ‘big ideas’ in their courses. In these maps, these ‘big ideas’ were referred to as ‘mountains’ with related basic skills residing in adjacent valleys (upper left) or in two dimensional rendering (upper right). In this case, a given student’s progress was described using gestures and graphics similar to those used in historical atlases – which describe phenomena such as the spread of a culture throughout some geographic region (lower left). As the maps show, some portions were mastered (lower right); others could be visited but were not yet mastered; and, still others were unfamiliar.”

3.1 Crucial Question: Not Just What Can Be Represented But How Much

There have been other attempts to sketch out what might be involved in creating visual displays of learner progress. Hoeft et al., and Wandersee [37, 38] discuss the adaptation of the growing field of concept maps with cognitive maps. Andrew Porter, former President of the American Education Research Association, also used the occasion of his presidential address to share some exploratory ideas relative to visualizations of learning [39]. The question posed here departs a bit from these types of efforts, not asking what
learning phenomenon can be graphically represented, but perhaps a larger and more strategic question: how much can be represented visually so that it can be meaningfully interpreted in one view? Why ask the question? While the state of the art of assessment practice is mixed, the issues are not so much that assessment tools are inherently deficient, but rather they are insufficient. In part, the focus discussed earlier on developing balanced portfolio of assessment information attends to this problem. But portfolios are challenging and time-consuming to construct, manage and interpret for students. The reason for asking the how much question here is that if the answer is “a lot,” that relatively large collections of visually presented information can be meaningfully processed by professors in a short period of time; an important new frontier for assessment in education opens.

To see this potential, it might be helpful to step out of the education research arena and reflect on other spheres of life where visual systems self-evidently pack large stores of information into graphical displays that are easily processed. Examples of what might be involved include the everyday weather report. Weather maps display huge amounts of information in ways that highlight relationships between storm or frontal subsystems and their animation accessibly depicts complex dynamic processes. Another example involves computer graphics in sports. Major League Baseball employs a website for disseminating play-by-play information as a game unfolds. Figure 2 is the display for one batter from the 2006 World Series. Hundreds of discrete pieces of information appear in the display. Several features are at work: observers from novice to expert can use the display to understand the flow of action in the game, though they may use the information in different ways. The interface is intuitive yet organizes a great deal of complex information. It highlights relationships between data, situating the current event (pitch location, ball in play, etc.) in the context of previous at-bats, previous pitches, the overall game, and the other current conditions in the game. The data displayed is heterogeneous, and the information is both formative and summative. Multiple temporal

Figure 2. “Enhanced GameDay” on mlb.com
This is the kind of data-rich and intuitive display that illustrates a vision for displaying large amounts of information that model learner classroom dynamics.
frames (speed of a ball; inning of the game) appear. For the present purpose what is most salient about the display is how much rich information can be displayed and processed by an individual in a short period of time. Such sophisticated graphics convey complex and nuanced relationships and information that can be processed very quickly.

Would it be useful or valuable for educators in the future to have the same level of rich, interconnected information about their students as Major League Baseball gives to its fans? What might be involved in displaying this much data and relational information about an individual’s or a classroom’s learning progress? This is an extensive design question, and one of the goals of this contribution to the workshop is to pose the first question in this paragraph and then to prompt informed discussion of how such displays might be structured.

One immediate visual element would be the display of what the student is doing in real-time, if the classroom configuration allowed transmission of workspace screen shots (such as with SynchronEyes or DyKnow). For example, the most salient part of the mlb.com display is the image of the batter. It is now possible to place on a teacher’s display a continuously updated image of the students’ workspace. That is raw, formative information, in the sense that it simply relates ongoing progress. More heterogeneous data sets can include information produced by instruments already in use such as recent course test scores, results from rubrics applied to a group activity, and learning progress maps. If the course included the use of intelligent tools or tutors, it could depict progress in a virtual world or other simulation, relate the frequency of help-seeking behavior or other self-regulatory activities, iterations along an ontology, text messaging in peer tutoring relationships, queries, and engagement data. (These are all predictable elements of future learning environments, and correspond to the principles appearing in Table 1.) These disparate types of data can be combined with historical data from traditional tests or other assessments, and can also include qualitative data. The goal is to produce graphics systems that process such data and display it in a structured way that is intuitive, rich, nuanced, and understandable. The systems should highlight relationships and furnish the basis for authentic data-driven instructional or other intervention decisions. The LPMs discussed earlier are a motivating example of potentially intuitive and information-rich displays that facilitate nuanced instructional judgements in short time frames and support a view of cognition that corresponds to nurturing problem-solving and modeling competencies. One reasonable expectation is that visualization systems like LPMs will help create new real-time sightlines into learner cognition and provide far more information than professors or learners are currently accustomed to processing in learning contexts.
The growth of visualization in educational tools has been deeply researched generating a significant literature on the cognitive advantages that emerge when richly visual representational systems convey course material. The suggestion here is to explore how much advantage might be obtained for the professoriate relative to assessment systems, especially as they become proficient or expert in efficiently encoding large stores of interconnected visual data. Such systems will inevitably entail new forms of expertise in encoding deep structure in a short period of time: a surgeon can read complex medical images and make rapid and accurate judgements about a medical condition because s/he has developed systems-level expertise. A fighter pilot makes split second decisions based on dozens of instrument readings and sightlines from the cockpit. In these instances, the professional does not and cannot use every piece of information, but through his or her expertise knows which information to process and what system relationships create the unique context that requires a decision. The encomium, “Experts not only do things differently, they see things differently [40]” is especially appropriate.

In each of these situations, sophisticated representational systems scaffold decision-making expertise. Such systems do not exist for educators, despite the plethora of potential data sources that could be combined. In fact, the use of such tools for conveying complex information about student learners is virtually non-existent. Websites such as http://www.visualcomplexity.com/vc/ (Figure 3, organized by information architect Manuel Lima) or the University of Maryland’s HCIL Laboratory http://www.cs.umd.edu/hcil/research/visualization.shtml are popular sites for taxonomizing or summarizing graphical modeling systems; they give examples of a broad cross-section of social, economic and behavioral processes but list none from education or for the assessment of learning.

4. Work Underway

The first principle identified in Table 1 for future learning environments centers on producing greater sightlines in classrooms. Visual systems that produce different ways to look at classroom learning are specifically intended to open new sightlines for the professor. Two projects carried out at the Air Force Academy and with partners at SRI and Boulder Learning Technologies focus on learning platforms that actually produce greater sightlines during instruction. One project, supported by Microsoft Research [41], explores the use of collaborative workspaces that network computers and permit exactly the type of real-time formative assessment displays that can thumbnail or expand the view of a student’s work in physics, engineering and mathematics courses. Another project, supported by NSF’s Computer Science Directorate, was one of the original motivators for the AGILEeViz exploration. That project, Agent and Library Augmented Shared Knowledge Areas (ALASKA) [42], integrates pedagogical agents, tablet computers, collaborative workspaces and digital libraries in STEM classes. Among other features, ALASKA creates a layer of classroom dynamics related to peer tutoring, asking questions of pedagogical agents, and accessing of digital objects. One intent of ALASKA is to relieve cognitive load on the teacher with this type of scaffolding [43], but in initial testing, the effects were just the opposite. Simply monitoring the classroom dynamics on various computer screens created an overwhelming burden on the instructor, and led to discussions of how to organize fast moving data in a visually intuitive way, in a way that highlights important relationships (such as how often a student seeks peer assistance, how often does the student give peer assistance, the kinds of questions the student asks an artificial agent, etc.) This quest to develop the visual systems is ongoing, and the project

is now engaged, through partners at SRI, with testing different visual mock-ups with expert teachers.

The types of visualization systems envisioned in this paper will require cross-cutting expertise and perspectives from fields as diverse as graphics design, psychometrics, eportfolios, software design, networking, pedagogy, and learner assessment. The hope of this paper is to engage multidisciplinary learning technologists and education researchers in conversations about the possibilities not of replacing assessment tools but of integrating the heterogeneous data forms they produce into complex visual structures that will facilitate far more informed and expert judgments and decisions about students in real-time.

Acknowledgments

Portions of the research presented in this article have been supported by Microsoft Research and by National Science Foundation awards 0433373, 0456434 and 0623166. The views in this paper do not reflect those of either organization.

This paper is a draft of a paper currently in review for the 2007 ICCE conference in Hiroshima, Japan and is provided as a discussion document for the July 2007 AGILeViz workshops.

References


41. Hamilton, E., T. Yamazaki, and L. Myers, *Comparative Evaluation of Four Modes of TabletPC

Use in Undergraduate Education. Grant from Microsoft Research, Inc., 2006.

Visualizing Student Learning and Understanding in Introductory Physics

Troy L. HACKER
Department of Physics, United States Air Force Academy
2354 Fairchild Dr, Ste 2A31, USAF Academy, CO, 80840
troy.hacker@usafa.edu

Abstract. Now that the effectiveness of active, student-centered teaching methods in college introductory physics courses has been well documented, instructors are trying to identify the best tools and techniques to use in class. The challenge is developing accurate measures to assess their impact on student learning and understanding. Performance on traditional assignments like exams and homework is a good start but these scores often create an incomplete picture of student thinking that can lead to misguided or incorrect conclusions. While developing new assessment measures is one option, this paper proposes a promising alternative that integrates many information sources including exams and homework into a single-screen graphical interface. Such a tool would allow physics instructors to visualize what students are learning in a format that is straightforward, intuitive, and meaningful. It would accomplish this by combining results from different activities given before, during, and after class into a visual display that captures any changes in the level of student understanding. Instructors could examine a conceptual breakdown to determine the value of a new teaching method or exercise, or they could analyze individual performance to identify and help students who are struggling. This application of visualization techniques could represent a significant advancement in assessing and improving student learning in introductory physics.

Introduction

In recent years there has been a concentrated effort throughout the physics education community to shift from lecture-based instruction to more interactive teaching that focuses not on what the teacher teaches but on what the student learns [1]. Instructors are now placing a greater emphasis on getting students actively engaged in learning the subject matter and exploring its applications and relevance through various activities in and out of class. But how do we know what students are actually learning? A first attempt at answering this question would involve evaluating student performance on traditional physics assignments: exams and homework. Depending on the method used, the results can offer a great deal of insight into what students are learning [2]. Unfortunately, deciding how to interpret a student’s score on a homework or exam problem to assess student learning is challenging at the very least, can lead to misguided or incorrect conclusions, and takes years of experience to master. Now that we are moving away from lecture-based, instructor-centered teaching toward interactive, student-centered learning, we must develop better tools for instructors which will help them answer the questions of what are their students learning and what can they do to better facilitate and guide that process.

Measures such as homework and exam scores are a good first step to gauge student learning, but they do not necessarily reflect student understanding of the subject matter in an objective, unbiased, and comprehensive manner. Taken individually they can
demonstrate a student’s ability to apply one or more physics concepts to solve a problem. However, if one integrates these scores with additional information such as GPA, scores on placement tests, and other demographics, one can begin to truly paint a more complete and accurate picture of what students are learning about physics. A particularly promising option is to create this picture in a single-screen graphical interface that is intuitive, user-friendly, and highly adaptable to the instructor’s needs. Such a visualization tool would have to integrate myriad sources of data to facilitate informed and objective decisions about the most effective tools and techniques for teaching physics. With this information one can quickly recognize topics that are particularly difficult for the majority of students, determine the value of a new teaching strategy, and identify individuals who are struggling early enough to give them tailored and timely instructional assistance. There are many other powerful possibilities for such a tool that helps instructors visualize their students’ performance. Ultimately, this can help transform the physics classroom into a dynamic learning environment that fosters curiosity and discovery about the world around us.

1. Assessing Student Understanding of Physics

With many instructors of college introductory physics courses now focusing on what a student learns rather than what is taught, we as a community must reconsider how we assess student understanding. This requires painting a whole-person picture of each of our students that goes beyond their scores on assignments. Even before the first day of class there are many pieces of information an instructor can use to gain insight into the strengths and weaknesses students bring with them into class. Some of these measures include GPA, SAT and ACT scores as well as scores on placement tests, high school records including AP courses, and so on. The trap that must be avoided here is automatically assuming students will perform at the level indicated by this information. These measures only begin to help an instructor recognize the group and individual needs of his students. Once classes begin student performance on traditional assignments such as exams and homework can help fill in the gaps and allow an instructor to tailor his lesson plans and one-on-one teaching to give his students the greatest chance of success.

1.1 Exams

For most college physics courses testing consists of 2-3 midterms and a comprehensive final that comprise the majority of the final grade. The exams typically include a mixture of multiple choice, short answer, and workout questions that test a student’s ability to apply physics concepts through problem solving. Some tests have been standardized so that comparisons can be made between different institutions to help identify best practices for the physics education community.

1.1.1 Traditional Exams

When it comes to studying for and taking an exam, students who are successful generally use one of two techniques: memorization or authentic learning. In many physics classes, instructors place too much emphasis on choosing the right equation and sticking the given numbers in the correct spot to get a solution. With this approach a student can easily memorize the applicable equations and how to use them and turn this knowledge into
high test scores based on getting the correct answer but not on demonstrating a comprehension of physics concepts. The undesired result here is this student leaves an introductory physics course with a good grade but essentially little to no insight into what physics means or how to use it to explain physical phenomena.

In order to shift from a computational and equation-driven approach to a focus on conceptual understanding, an instructor needs to put much time and thought into the types and format of the exam questions, the desired skills that a student needs to demonstrate, and how the exams are graded to accurately and fairly assess what students are learning. How to accomplish this task is beyond the scope of this paper but, when done properly, scores on traditional exams can provide useful and relevant measures of student understanding [3].

1.1.2 Conceptual Exams

Since exams vary widely across colleges and universities, they are not very useful in assessing student understanding on a larger scale. We can try to address this issue by administering standardized exams that departments can use to benchmark and compare student performance. For introductory physics courses, the Force Concept Inventory (FCI) [4] and the Conceptual Survey in Electricity and Magnetism (CSEM) [5] were developed for this purpose. These tests are typically administered at the beginning and end of the course to measure the changes in student understanding of physics concepts after a course of instruction. Scores on the FCI and CSEM can give instructors reproducible and consistent results about how their students compare to those in the past at their own institution as well as other schools that also administer these exams.

1.2 Homework

After exams, homework is the next highest assignment score in a student’s final grade. The main purpose of homework is to give students the opportunity to practice the skills and knowledge they will be required to demonstrate on exams. Usually an instructor assigns a selection of problems from various resources including textbooks and personal or published question archives. Until recently homework was done by hand which made grading a challenging and time-consuming process. There are now numerous online physics homework systems that allow an instructor to select homework problems for their students to do online and submit their work electronically which is graded automatically. While this addresses the instructor burden associated with written homework, there are other problems introduced with online homework that must be considered when using it as a tool to assess student understanding of physics.

1.2.1 Written Homework

When a student turns in a hand-written solution to a homework problem, the potential insight gained by the instructor into the student’s knowledge and understanding goes way beyond the answer. Now one can see the problem solving method used to go from the problem statement to the answer. This perspective is perhaps the best measure of authentic learning but takes a significant amount of time and effort for many reasons. First, following a student’s solution that may contain mistakes, omitted steps, and poor handwriting makes grading a very time-consuming process. More importantly, an instructor’s ability to recognize the desired level of student understanding from homework is extremely challenging because he is making a judgment about the student’s thought
process based on a written solution. To do this properly one must put aside personal
preconceptions and look at the problem from the student’s point of view. This skill takes
many years of experience for instructors to develop which can sometimes be a detriment
to students. In the end, written homework can be an especially effective tool for assessing
student understanding but not without much trial and error along the way.

1.2.2 Traditional Online Homework Systems
In recent years, advances in computer and networking technology have spearheaded the
development of online homework systems for numerous age groups and subjects. While
there are multiple products available for introductory physics including PhysicsNow,
WileyPlus, ARIS, iSolve, and WebAssign, my experience is primarily with
MasteringPhysics so I will focus on this system. It has many strengths including
automated grading with a convenient gradebook that tracks student homework scores and
time spent on each problem and displays this information in a well-organized, easy-to-
read format that an instructor can use to see how his students are doing on their
homework assignments. Furthermore, MasteringPhysics contains an extensive problem
library with many different types of problems such as ranking tasks, graphing, and
drawing vectors that have been shown to force students to confront their misconceptions
and develop a better understanding of physics. Many problems include in-depth
explanations with a series of hints that students can use if they get stuck when trying to
solve a problem. An instructor can choose a mixture of skill builder, self-tutoring, and
traditional end-of-chapter textbook problems that best achieves desired learning outcomes [6].

In my experience, the convenience and potential benefits of MasteringPhysics for
me as an instructor do not necessarily translate to improved student understanding of
physics. In fact, online homework poses new obstacles to learning that an instructor must
address. Aside from the fact that it is much easier for copied answers to go unnoticed in
the absence of written solutions, online homework assignments do not give an instructor
insight into how a student solves a problem. The skill builder and self-tutoring problems
in MasteringPhysics are designed to help students develop a conceptual understanding of
a particular physics topic and see how it is used to solve a problem or make a conclusion
about how the world works [6]. Then they to apply this knowledge to end-of-chapter
textbook or similar type problems where they can no longer rely on detailed explanations,
multiple steps, and hints to arrive at the solution. While this is the goal many students are
not able to reach it simply by doing homework on their own, and if they do not ask for
help an instructor has no way of knowing who is struggling until they do poorly on an
exam. Online homework unfortunately comes up short in giving an accurate and complete
assessment of student understanding so an instructor is forced to find alternative methods
to evaluate how students solve physics problems.

1.2.3 An Artificially Intelligent Online Homework System
Despite all the features designed to help students better learn physics, nearly every online
homework system currently available is basically programmed to check a student’s
answer against an approved solution. We need to find some way to combine the
convenience of online homework with the insight into thought processes gained from
written homework. I think this has the potential of providing the greatest benefit to both
instructors and students and can be achieved by incorporating an artificially intelligent
agent into an online homework system. This agent essentially serves as a personal tutor that enables a student to learn and apply physics concepts by providing tailored help when needed. These hints are based on evaluating the student’s work to that point, determining what is missing, and guiding him to the point where he is able to figure out what to do next without explicitly telling him. Now this would never replace an instructor but could possibly be a powerful tool to maximize student learning of physics.

An artificially intelligent online homework system with these capabilities will take many years to create, but currently there exists a system that represents a first step on the road to this lofty goal. The system is called Andes and was developed at the University of Pittsburgh and the United States Naval Academy [7]. Andes goes beyond having students enter an answer and requires them to enter all the steps they use to arrive at a solution including drawing coordinate systems and vectors, defining variables, and writing out equations in symbolic form. An artificially intelligent solver checks each entry and provides immediate feedback by turning it green if correct and red if incorrect. The solver also includes a help function that gives principle-based hints when requested based on all the entries to that point. This step-by-step problem solving method required by Andes mirrors the approach we encourage our students to use when doing these same problems by hand, but the feedback and help features have been shown to reduce student frustration, minimize the time required to do homework, and increase learning [7]. Andes does have some drawbacks including a comparatively small problem library, a cumbersome workspace with a steep learning curve, and limited technical support. However, after personally using Andes for two semesters I think it has great potential to revolutionize student learning in physics by using an artificially intelligent agent to guide students through a wide variety of homework problems.

2. Visualizing Student Understanding of Physics

Regardless of the subject matter the primary role of an instructor is to help students learn. If this authentic learning is our principal goal, then the question we must ask ourselves is how do we know if what we are doing in the classroom is achieving this result? It is so easy to limit our self-assessment to final grades and student course critiques and use these measures to draw comparisons between different instructors and attempt to make conclusions about successful teaching techniques. The harder result to obtain but arguably one of the best measures of instructor effectiveness is to find out what students still remember from our courses five and ten years down the road because this will tell us what was truly learned instead of memorized and forgotten. Such a far reaching evaluation is not very helpful in determining the best approach to help students learn a new subject now. So how can an instructor assess student understanding during the course of a semester to maximize learning? For many of us this is simply a skill that we must cultivate with years of practice, but it does not have to be this way. I think that a visualization tool that consolidates numerous sources of information into an intuitive and meaningful graphical interface could be invaluable in helping instructors gauge what their students are learning so that they can make worthwhile adjustments to what and how they teach.
2.1 Graphical Interface for Instructors

The most obvious questions concerning the creation of a visualization tool for instructors is what information should be included and how should that information be displayed? We can start by looking at other situations where such a visualization tool is used effectively. In my opinion, the major sports leagues really have this idea figured out. One can go to the websites for any professional sport, click on a link for a game in progress, and “watch” the game real-time on a single screen that is visually appealing and easy to follow. This screen continually updates with live statistics and play-by-play results such as the 95-mph fastball that caught the outside corner for a strikeout to end the third inning in a baseball game or the handoff to the running back up the middle for a 12-yard gain and a first down on the 45-yard line in a football game. These tools consolidate and display an enormous amount of data in a format that is virtually the same as watching the game live or on television. All a fan needs is a basic understanding of the game to use these tools to keep up on his favorite sports teams. In education this type of visualization tool would allow an instructor to quickly and accurately create a real-time, big-picture assessment of the level of student understanding throughout the course of the semester. For me personally, my overall teaching goal is to have my students leave their physics course with a conceptual understanding of how to use physics to explain natural phenomena and the ability to frame and solve complex, ill-defined, and technical problems. Such a powerful tool could help me determine if I am achieving this goal in a way that would ultimately help me be a better teacher for my students.

2.1.1 Format and Organization

A visualization tool for an introductory physics course must organize and display relevant information to enable an instructor to make a quick, useful, and straightforward appraisal of student performance with respect to desired learning outcomes. It must be able to access multiple, separate databases and integrate everything into a logical and user-friendly format on a single screen. Additionally, the tool should allow an instructor to tailor the information and displays to meet his individual needs and preferences.

A rough draft of the format and organization of this visualization tool is illustrated in Figure 1. When first accessing the tool, I would want to see a reflection of the current progress of all my students in the course. This would be organized by conceptual topics common to most introductory physics course with a graphical display such as a histogram. For each topic, the bar height would reflect the level of individual and collective student understanding as demonstrated on tests, homework, and other course activities. As topics are introduced, explored, practiced, and integrated during the course of the semester, the bar heights would ideally increase. I would also want to click on specific topics or students and see customized displays that will help me evaluate the success of various course activities at facilitating desired learning outcomes (shown in Figure 1) or identify the source of individual student difficulties. Overall, the visualization tool must have a format and organization that is easy to follow, flexible, and comprehensive if I am to use it to make informed and constructive decisions about how I teach physics to my students.
2.1.2 Information Resources

As described in Sections 1.1 and 1.2, some of the information included in a visualization tool for physics instructors can and should still come from traditional exams and homework assignments. The challenge is to figure out the most insightful and meaningful way to display this information. Raw homework and exam scores are not very good indicators of student learning by themselves. However, if they are combined with other information such as applicable concepts and equations for a given problem, the approach used to solve it, the time spent to complete it, the mistakes made along the way, and the amount and type of help requested, the resulting picture would be enormously useful to determine what students are actually learning. For exams and written homework, much of this information would have to be entered in by hand. With the exception of time spent completing a problem and hints accessed, MasteringPhysics does not capture this information either [6]. Only Andes is currently capable of providing objective numbers that reflect hints needed, correct and incorrect entries, and the drawings and equations used to solve problems [7]. This information shows how students think about and apply physics concepts in a unique manner which would be invaluable in a visualization tool that assesses student understanding.

Beyond exams and homework, there are numerous other resources this tool should access and display that are useful to instructors. Along with student performance on class assignments, additional personal information such as GPA, academic background and
major, standardized test scores (e.g. SAT, ACT, AP, and placement tests), class attendance, and extracurricular activities would all give insight into students’ lives outside class that could affect their effort and performance. As for the graphical topic display described in Section 2.1.1, the information used to determine the bar height should change based on current progress in the course. When a topic is first introduced, the height should be relatively low based on standardized pre-test scores, warmup exercises before class, and student reading. Once the topic is covered in class, instructor and student subjective assessments of the lesson as well as performance on classroom activities like quizzes and clicker questions [8] can be factored into the bar height. Finally, exams and homework that allow students to practice and apply the topic further adjust the height. From start to finish this collection of numerous and diverse pieces of information will give an instructor a more complete and unique measure of student learning based on the current and historical bar heights.

2.2 Teaching Aid for Instructors

A visualization tool that helps instructors assess student understanding is futile unless it is used to influence classroom activities and teaching methods thereby helping students achieve their fullest learning potential. Most of us are willing to draw upon the good ideas and best practices of our colleagues to give our students a first-class physics educational experience. But how do we know which tools and techniques are most effective to learn physics, and when do we make large-scale changes affecting the entire course versus smaller ones targeted at individual students? Referring to Sections 1.1 and 1.2, we currently use measures like exams and homework scores, standardized tests, and final grades to attempt to answer these questions. However, the visualization tool described in Section 2.1 could revolutionize the approach we use to compare and select different teaching methods and determine their respective merits and shortcomings.

2.2.1 Course-Wide Evaluations

In an introductory physics course, the students come from a wide range of backgrounds. This is especially true at the United States Air Force Academy where all cadets take two semesters of physics as part of the core curriculum. Teaching methods that work for a history major are not necessarily as successful for an engineering major and vice versa. These methods must continually evolve based on individual student performance and advances in physics education. Our ability to choose the best method for a particular topic and group of students is one of the primary factors that influences their ability and desire to learn. Instead of relying on imprecise and flawed metrics like exam and homework scores alone, a comprehensive visualization tool can help us identify the most effective teaching methods and design course activities that maximize learning for our students. It can also highlight areas that are particularly challenging for most of the students and, therefore, require additional attention in class.

For example, suppose an instructor wants to try a new classroom activity to clarify this difficult topic. Afterwards, he can use the visualization tool and see if there is any increase in the bar height associated with that topic. The new activity may have helped one student achieve a better conceptual understanding of the topic while another may have adapted it to a homework solution. Traditional metrics may have captured the latter student’s improvement in terms of a higher homework score, but a comprehensive graphical representation as shown in Figure 1 would ideally recognize all possible effects

25
on student learning. The potential of such a visualization tool to enable big-picture evaluations of the overall effectiveness of various teaching tools and techniques is unmatched by anything that currently exists and promises to give the greatest benefit to the biggest number of students.

2.2.2 Individual Student Performance

Beyond making informed decisions about the best teaching methods to use in class with everyone, the ability and desire to learn physics varies with each student. When a student fails an exam, there could be many reasons that explain why. Perhaps he is dealing with personal problems and is doing poorly in all his classes as a result. Maybe he struggles with math and other technical subjects and believes he does not have the talent to succeed in physics. It could be that he simply did not study enough and does not really care about his performance. Although talking with this student is the best option to figure out what went wrong, a visualization tool would help an instructor narrow down the possibilities and come up with ideas for improvement that best draw on the student's talents and interests. Such thoughtful support would go a long way in proving to our students that we really do care about them personally and want to do everything we can to help them learn and grow.

3. Conclusion

It is now time to expand our view beyond exam and homework scores and create a comprehensive and insightful visualization tool to allow instructors to truly find out what their students are learning in their introductory physics class. Furthermore, such a tool adapted for student use that allows them to visualize their individual progress in a way that is simply not possible with assignment scores and grades would also have significant implications for learning. Thousands of educators and researchers nationwide are trying to develop effective teaching tools and techniques so that everyone has access to and can benefit from a physics education. As mentioned previously, our ability to definitively conclude that a particular teaching method leads to increased student learning is difficult to say the least and will never be an exact science. We must, however, come up with something that we can all agree upon that goes beyond traditional measures like exam and homework scores. I think that a concept-based, user-friendly, and highly flexible visualization tool could bring about powerful new insights into the learning, understanding, and application of physics knowledge that are just not possible today.

References


Making Sense of Complex Learner Data

Christopher Brooks, Wengang Liu, Collene Hansen, Gordon McCalla, Jim Greer
Laboratory for Advanced Research in Intelligent Educational Systems (ARIES)
Department of Computer Science, University of Saskatchewan, Canada

Abstract. Over the past decade, we have been involved in building and deploying a number of e-learning applications with the expressed goal of collecting detailed usage information so that better personalized systems could be developed. This paper outlines our most recent research in the collection, processing, and display of these learner traces.

Introduction

Modern higher education e-learning environments are made up of a plethora of different tools for student interaction with both content and with their peers. Routinely, instructors are using content management systems (e.g. WebCT, Blackboard, Moodle, etc.) which typically include the ability to deliver web-based content, evaluate student performance from web-based quizzes and interactive assignments, and facilitate discussion through synchronous (chat), asynchronous (forums) and semi-synchronous (instant messenger) means. These environments are being used both to supplement traditional education approaches (e.g. a blended learning approach), as well as to facilitate complete online and distance learning.

Over the past decade we have been involved in the development and deployment of a variety of e-learning tools (e.g. [3], [4]). One of our key concerns while developing these tools has been to include in them the ability to create fine-grained user models based on usage information. These models can then be processed for a particular purpose and used as a reflection of student achievement for both instructors and learners [5]. We see the presentation of learner models, both individual student models and models describing cohorts of students, to be of increasing importance as more e-learning tools are included in course curricula. We are currently undertaking three separate but related approaches to dealing with the processing and display of complex learner data.

1. A Data Mining Data Model for Usage Information

Raw usage data is generally too fine grained to be of any use directly to an instructor or student, and needs to be processed into larger situation-dependant “chunks” instead. We have been using a number of manual and automatic methods to create this higher level information. One of the methods we are employing (Figure 1) is the use of automatic data mining techniques, where the partially processed raw data is aggregated into slightly larger pieces of information (which we call facts) and then processed using a variety of clustering algorithms. These algorithms also apply a series of instructor-defined measurements that can be combined into metrics related to pedagogical theories or goals. The output of this process is a ranking or grouping of learners based on the metric being investigated, along with a label which described that metric.

Currently we are using the EM, K-Means, and X-Means clustering algorithms through the Weka toolkit.
In an upcoming pilot study of this approach, we are going to investigate four metrics, each describing a different abstract characteristic of learners: their level of activity and interactivity, their level of socialization, their learning style, and their knowledge level. Each metric is made up of a number of measurements based on the instructor’s professional experience as well as any pedagogical theories they might employ. For our pilot study, we have identified the following relevant measurements:

- **Interactivity**: the course content pages the student has visited, the asynchronous messages they have read and written, the synchronous chats they have been involved in, and the quizzes that they have finished.
- **Socialization**: The presence, responses, and connectedness of the learners in the environments based on [1].
- **Learning Style**: Whether the learners are global or sequential learners, based on [2]
- **Knowledge**: The correctness of learner responses to content-based quizzes, and time spent on pages.

In the study, there will be no predefined classification or category for each measurement. Each time instructors query a specific measurement, a clustering algorithm will compute and create new groups based on current fact data distribution. In different stages of a study term, measurements will be dynamically created, and learners will get different labels if their behaviours change throughout this period or if fact data distributions change. To evaluate that these metrics and measurements are useful for instructors, we have designed a human evaluation procedure to judge whether or not experts share similar judgements.

### 2. Real-Time Reflective Modelling

For online and blended learning courses, it is difficult to measure how a student is performing, and whether external actions must be taken to encourage behavioural changes to improve performance. Online delivery systems capture a large amount of raw data, but the data alone provides no help without an established context (meaning) behind the data. We have developed an online query tool an instructor can use to create queries with purpose to ask particular questions of the online delivery system. For example, the instructor may query the system to ask “who is falling behind?”. Using this information, the instructor may then communicate these concerns to the students to challenge them to
do better.

One observation early in this project was that different instructors have different opinions on what it means to “do well” in a course, making computation difficult for every possible situation. Instead of forcing a system-specific calculation and having instructors adapt, instructors can build queries themselves appropriate to their purposes using raw data and analysis instructions. They can then run the queries to view the results. The information is computed in real time when selected, so the data is always up-to-date. Figure 2 shows a query made by an instructor to determine who was falling behind in a blended learning course. In this case, the instructor defined falling behind as students who had not viewed the appropriate lecture videos. When run, the query shows 9 students were falling behind (identifying information has been grayed out for privacy).

Instructors may also choose to allow students to run the query, enabling student reflection on their progress in the course. An example of a student-enabled query is shown in Figure 3. Privacy protection is in place to block identifying information such as names from peers, but still enough information is displayed to be useful to the viewer. Students are able to see all the information about themselves and the viewer of the query is highlighted in yellow on the page for quick comparison.

Figure 2. A purpose created by an instructor (left) and the query when displayed (right).
A formal study is currently underway to analyze the effectiveness of the open learner modelling tools in several courses. However, informal communication with instructors and students has already indicated an appreciation for the system. From one student in an online course: “Thanks for putting up this compare my progress information. It is really interesting to see how we are doing and helpful. I like to be able to see my progress, it makes me feel like I have accomplished something.” Student behaviour has also changed. By seeing what was expected for higher participation marks in the course and by having the ability to run the query throughout the course, students increased interaction through the discussion forums in attempts to raise their participation in the course.

3. Visualizing Metrics

Getting a view of the “big picture” of the social interactions between students becomes increasingly difficult as class sizes grow and face-to-face learners become distance learners. However, the understanding of the social relationships between learners (often referred to as social capital) is incredibly important when an educator is determining how to scaffold the learning content. As the amount of data increases, it becomes harder to get a view of the big picture in the classroom; however, focusing solely on the big picture often ignores the subtle details that convey why the big picture is as it is (that is, it is in the details that an instructor can often see the immediate effect of a particular pedagogical practice). We have been working on isolating community interaction using a set of nested sociograms, where each sociogram is made up of nodes that represent users (instructors and support staff are nodes which are coloured red), and edges represent a direct “reply-to” relationship within an asynchronous forum. We break users up into three different categories based upon their activity levels:

- Participants: Those individuals who have written messages, either on their own or as replies to other messages.
- Lurkers: Those individuals who have read postings but have not written any.
- Delinquents: Those individuals who have never read nor written a posting, but are in an institutionally defined role that gives them access to the postings (e.g. enrolled in a course).

Each class of users is put into their own sociogram with nodes aligned along the
exterior of a circle. The different sociograms are then layered on top of one another such that the delinquents are farthest from the centre of the screen, the participants are closest to the centre of the screen, and the lurkers are in between (Figure 4). This corresponds closely both with the perceived participation rate of individuals, as well as with the sizes of the different classes of individuals in most circumstances. We calculate the rate of lurking of users compared to one another, and vary their distance linearly from the outer edge of the lurker region to the inner edge based on the number of messages that have been read. This reinforces the notion that users who are close to the centre of the screen are participating more in the course than learners closer to the edge of the screen. We use node size to differentiate the level of perceived importance that a particular participant has in the community. We define importance as the effect that a particular person has on the immediate reading habits of others (e.g. the number of a person’s postings which are read). The values of importance are then mapped to nine different node sizes. Everyone having an importance value of less than 0.5 is mapped to the smallest node size, and everyone having an importance value greater than 3.5 is mapped to the largest nodes.

![Figure 4. A cut away portion of the community visualization tool.](image)

In Figure 4 the outer ring captures the set of delinquents, the inner ring captures the lurkers, and the middle circle shows the relationships between posters. Node 1 shows a particularly keen student, who has both initiated and responded to many postings, and whose postings have been read by many peers. Node 2 shows an instructor – the large number of edges indicates that this instructor has responded to many postings. Finally, Node 3 shows a lurker who has read more than half of the messages (e.g. they are closer to the participants than the delinquents). The numerous red dots in the delinquent ring identify that a large number of tutorial assistants (general paid helpers) and faculty have access to this class forum, as is common at our institution. Initial response to these visualizations from instructors has been good. We are currently undergoing a study looking at how specific pedagogical practices (such as scaffolding conversations during debates or role playing) can be realised through these kinds of visualizations.

4. Conclusions

Anecdotal evidence from both instructors and students has suggested to us that being able to visualize and summarize the kinds of data being collected in e-learning systems is of high importance. In this paper we have outlined three complementary projects that are looking at issues involving the collection and summarization of user model data, control
over custom reporting of such data, and the visual display of such data in pedagogically informative ways. The common threads weaving through these three projects are, first, the importance of giving users (pedagogues and students) control over the kinds of feedback they want from an e-learning system; and, second, basing this feedback on real data about student performance gathered by the e-learning system itself. Formal studies for each of these projects are currently underway.

Acknowledgments

These projects are funded in part by the Government of Canada through the LORNET NSERC research network.

References

The Social Affordance of a Paper

Tiffany Y. Tang and Gordon McCalla*
Department of Computing, the Hong Kong Polytechnic University, Hong Kong
*Department of Computer Science, University of Saskatchewan, Saskatoon, Canada

Abstract. In the past we have proposed a set of recommendation mechanisms that are suitable for recommending papers to learners who are learning a new area of research. In this workshop paper, we shift our attention to the design of a recommendation system that draws on social affordances in supporting recommendation and social interaction in a learning environment. In particular, through displaying the social affordance of a paper, the paper recommender can make learners' learning activities, learner interaction more efficient.

Introduction

Research in visualizing students’ activities in computer supported collaborative learning (CSCL) environments has focused on monitoring activities in discussion, performance on quizzes, and content access patterns [1]. Reffay and Chanier have implemented a Distance Learning Management System that can identify clusters and cliques among interactions [2]. They focus on the activity of the class as a whole, not on the individual student, as is emphasized in Hiltz and Turoff [3]. Counting the number of messages posted by a student is one of the most critical and reliable ways to measure student participation in an online environment. However, as Saltz, Hiltz, and Turoff have pointed out when we are counting the number of messages posted and authored by a particular student as one of the performance indicators, care should be taken to avoid “superficial posts” [4, 5]. These so-called superficial posts might have low value, which results in few follow-up messages. In other words, when the number of follow-up messages increases, it is safe to say that the student has had some impact on the whole class [4]. This notion reflects the social reinforcement relationship in an online distributed environment.

In CSCL, the students, tutors and learning materials as well as the interactions among these entities are all very important to draw a complete and coherent picture of the learning as a social process. To achieve this, the system should allow information spaces that afford social interactions as much as possible. While the majority of existing systems tend to provide systems to support awareness to make coherent discussions in forums, the goal of our social tools is to make learners’ awareness of not only themselves in relation to others, but also the objects they are manipulating. In our research, we have explored this issue in the context of a paper recommender for collaborative learning environments [6, 7, 8]. In the paper recommender, the system picks up an article for learners based on aggregated observations of each paper’s pedagogical values, such as its usefulness in helping learners gain new knowledge and strengthening their understanding of the course topics, its practical value, the degree of peer-recommendation, instead of purely basing judgement of a article on its interest to the learner. By dividing the paper information space into smaller, finer elements, learners can move freely between the papers, make multi-dimensional annotations and make decisions on picking appropriate papers, forming discussion groups, formulating
team projects etc. The confluence of factors that affect the usefulness of a paper to a learner is called the social affordance of the paper. We would like to make available this social affordance to the learner, that is to make visible the relevant aspects of social interaction about a paper. We would like to unfold various relevant aspects of a paper, and meanwhile, reflect the features of the learners who are interacting with the paper.

In this working paper, we overview the design of a recommender system that will draw on social affordances to recommend papers to learners in a learning environment. In the next section, we will review affordance and social affordance in general and in the particular domain of CSCL. We will then focus on the social affordance of papers in the educational setting of a paper recommender system. Discussion of related work and a synopsis of the state of our research conclude the paper.

1. Providing Social Affordances for each Paper in the Paper Recommender

1.1 Affordance and Social Affordance

The term, affordance, originated with psychologist James J. Gibson in 1977, and refers to all the ‘action possibilities’ latent in the environment, objectively measurable and always in relation to the user [9, 10]. The term has received much widespread popularity since 1988, when Donald Norman introduced and appropriated it in the human-machine interaction area [11], in which affordance encompasses all those action possibilities that are readily perceivable by an actor. As such, affordance depends not only on the physical abilities of the actor, but also their meta-cognitive skills such as planning [11].

Take a media player as an example to illustrate the term. The existence of the functional buttons on the media player affords users a visible cue that they can press a button, or roll a wheel just like as they do in the Apple iPod. If these buttons are well designed and displayed distinctively from the rest of the player (say, grouped in sections, and highlighted in different colours/shapes), the ‘affordance’ cue is emphasized, which, in turn, can enhance the ‘usability’ of the devices. In the human computer interaction area, affordance remains as one of the most essential elements to ensure high usability.

According to Kreijns, Kirschner, and Jochem [12], social affordance in a CSCL environment refers to any elements of the learning design and environment that would help facilitate students’ learning. They further argue that ‘a good set of social affordances will establish the desired sound social space that is characterized by an effective structure, trust and belonging...’ [12, p. 7]. Social affordance of an item in a CSCL environment can allow users to perceive their environments (which include the items, other users, as well as the system), formulate their actions and navigate in the information space accordingly. Making such social affordance visible would be useful and would keep social affordances at the heart of any CSCL environment.

1.2 From Social Filtering Supporting Collaborative Learning to Social Affordance

1.2.1 Social Affordance in a Social-Filtering-based Paper Recommender

Collaborative filtering (CF), also known as social filtering, is one of the major techniques for recommender systems. CF relies on ‘word-of-mouth’ social methods to make personalized recommendations. A CF-based system makes personalized recommendations to a target user mainly based on the ratings provided by other users as well as the target user him/herself. CF has been widely explored in such commercial systems as Amazon.com, CDNow, and others.
as a convenient as well as effective way to help thousands of users find the most useful information among millions of candidate items.

Obviously, the fundamental feature that differentiates collaborative filtering from other types of recommendation mechanisms is its ability to rely only on users’ numerical ratings (except for the critique-based conversational recommender systems [13]) which are independent of the content features of the system. When simply displaying the rating, though, the system fails to contextualize the popularity of the movie since users might want to see why the number is what it appears, as Lueg [14] argues in the problem of document recommendation, these numerical values might be de-contextualized:

For example, it is assumed that the “content” of a document can be observer-independently estimated on the basis of its representation. Also, it is assumed that “interest” can be estimated independently from the actual situation the recipient of information is involved in. Accordingly, it is assumed that ratings given by a particular person in a specific context can be appropriately represented in numeric ratings and that it makes sense to de-contextualize these ratings. [14, p. 2]

To compensate for de-contextualisation, in some commercial systems such as Amazon.com, the rating would usually be followed by editorial reviews and some readers’ comments that identify several situational factors to influence users’ acceptance and understanding of the numerical rating. Additionally, users might also be eager to see their neighbours’ feedback on the movie. So how to make the recommendation result fairly direct and transparent to users has recently become a focus [15, 16].

From the situated cognition perspective, the simplified rationalistic perspective underlying the majority of recommender systems does not appropriately capture human information-seeking and cognitive behaviours. One key implication of such situated-ness is that the way a user interacts with the system will continuously change based on his/her perceptions of the ‘environment’ where the interactions occur. Accordingly, in order to allow users to formulate comfortable and personalized context, the system should provide and unfold well-designed social affordance cues to facilitate social interactions. In other words, the concepts of social affordances can guide the design of ‘group-friendly’ social interaction tools or systems including large-scale commercial systems such as MySpace.com and Facebook which have gained wide-spread popularity recently.

To show the role of social affordance in a general recommender system on movies, let us review the following five possible ways of depicting the popularity of the movie “Notting Hill” as depicted in Figure 1.

The first three simply reported the ratings either numerically, or visually (as a bar indicator); while the third compares the acceptance of the movie with that of one of the most popular one in this genre, namely ‘Titanic’. Notice that the last one differs from the first four in that it simultaneously presents the ratings of the target users’ neighbours, which highlights the social affordance of the movie, because the user might want to know how other users with whom he/she shares similar tastes judges the movie (which is also fundamental to social recommendation, although the value again is de-contextualized). The last visualization method was studied in [17], and the experimental result shows that users, on the average, tend to favour it over other visualization alternatives, though, theoretically, it may not be feasible to do so due to the large quantity of neighbours, especially in commercial systems. One remedy to this is to pick some neighbours according to some criteria and show them only. The Herlocker study fails to analyze the impact of this design on supporting the social interactions among the users, and between the user and the system. In fact, in the recommender system literature in general, although visualization is essential in increasing user acceptance, very little research addresses this issue so far.
1.2.2 The Social Affordance of a Paper

Barry [18] pointed out that situational factors (contextual factors) other than only the topical content of a selected document influence a user’s judgment of document relevance as well as quality. He suggested that these situational factors include those that users bring into the reading situation including experience, background, knowledge level, beliefs, and personal preferences that should be added are the co-existence of other users’ traces on the document, including the social annotations such as ‘thumbs up’ ‘thumbs down’ in KnowledgeSea II [19], the textual comments, popularity of the each article and user models annotated to the document; in other words, the social affordance of the document. When users browse a digital document space, these elements can reveal the situational factors that could influence a document’s overall user acceptance.

In our previous work [6, 7, 8], we proposed a pedagogical paper recommender, where certain characteristics related to learners themselves were incorporated when making recommendations. Learners were asked, in the paper feedback form, to respond to a variety of Likert-scale questions regarding not only the overall rating of the paper, but also ratings pertaining to the paper’s ‘situated factors’ including its usefulness in helping learners gain new knowledge (referred to as Value_addedness) and strengthening their understandings of the course topics (referred to as Aid_learning), value in practice (referred to as Job-related, as learners were all part-time degree students), the degree of peer-recommendation (referred to as Peer-Rec), and textual comments\(^3\). Figures 2 and 3 compare our way of making recommendations with that used in the majority of recommender systems.

---
\(^3\) A deeper discussion of this is beyond the scope of this paper.
In Figure 2, the information space allows users to review both the textual comments and the numerical rating of a movie. However, the recommendation mechanism only considers the latter in addressing users’ needs. In contrast, our pedagogical paper recommender works by incorporating learners’ additional impressions of each paper. Figure 3 illustrates the recommendation mechanism.

**Figure 2.** An illustration of user relevance evaluation on a movie

Recommendation is based on overall rating only

In the pedagogical paper recommender, a paper is recommended based on a variety of dimensional factors that a learner has provided in terms of not only its overall rating of the topical appropriateness, but also some pedagogical values (situated factors). Our goal is to understand the many factors driving learners to judge the ‘goodness’ of a paper. When the
system allows users to unfold these aspects of a paper, it actually creates a rich space for learners to interact with the system and other learners. For instance, it can help raise the awareness of a learner towards the candidate papers or provide an opportunity for the learner to socialize with others through initiating discussions. Our experimental studies confirm our speculations that making recommendations to learners in social learning environments is not the same as that to users in Amazon.com etc. Learners are willing to accept those items that are not interesting, yet meet their learning goals in some way or another; learners’ overall impression towards each paper is not solely dependent on the interestingness of the paper, but also other factors, such as the degree that the paper that help to meet their ‘cognitive’ goals [8].

From a user-recommender interaction perspective, we should deliver the recommended items appropriately so as to show their pedagogical values, i.e. the social affordances of the paper. To further illustrate our ideas, Figure 4 lists some possible ways to manifest a paper’s social affordance which are aligned with the algorithmic design of our paper recommender. For example:

- Paper A not only has an overall rating, but also has ratings in other dimensions, including the user’s neighbours’ ratings.
- A popular paper tag map is capable of making the most well-accepted tag(s) visible.
- The paper’s position among a pool of similar ones can be shown.

These visualization alternatives can clearly reveal the social features of a paper in relation to other papers.

Our goal here is not only to visualize each paper’s standing in the system, but also to

![Figure 4 The Social Affordances of a Paper](image-url)
make users aware of many aspects of a paper. That is, one task is to investigate how one user’s behaviour and tastes can be influenced by seeing other users’ likes and dislikes. It is known that people like to watch what others are doing, a phenomenon known as ‘info-voyeurism’ by the founder of Plum Ventures, Hans Peter Brondomo. The company recently released a product called Plum (http://www.plum.com/), which allows users to collect a wide variety of information ranging from web sites and photos to music and text files, add comments and share the information with others. Although in the research community, especially the recommendation community, there is very little research on how a user can influence another user, we hypothesize that a user’s tastes might be changed when looking over another’s shoulder [18]; hence, it might not be surprising that there are gaps between users’ picks of papers and system-recommended items. However, our current focus in visualizing each paper’s social aspects is to raise awareness among learners by offering a highly flexible social affordance on each paper, and meanwhile by allowing the users to perceive this affordance when working with the document, and by allowing learners to be able to see what kind of learners like to read certain articles.

2. Related Work
There are two main branches of related work, i.e. social affordance and social interactions in Computer Supported Collaborative Learning and recommender systems, and visualization in CSCL which has been presented in Section 1.

2.1 Social Affordance and Recommender Systems

Although there are numerous studies documenting research methodologies and practices in recommender systems, very few focus on studying how to go beyond just the numerical ratings of an item could influence user choice. One that does is Herlocker, Konstan, and Riedl [17] which studies how the ratings should be presented to the users so as to not only inform them of the popularity of a given item, but also to implicitly influence their choice. Unfortunately, the paper fails to relate the approach to the issue of affordance. Jameson, Baldes, and Kleinbauer [20] argue that providing mutual awareness in the recommendation interface could help users arrive at a consensus when each of them has a chance to specify what they like and dislike in a group travel recommendation. They further pointed out that a mutual awareness environment can benefit its group members by avoiding their investing time on post-recommendation communication; that is, after the system makes recommendations, group members should be able to quickly arrive at a decision with much less discussions. In our approach, though, users can access the information on the artefacts to gain awareness of how the artefact is being viewed by different kinds of learners. By doing so, a learner can help him/herself to gain an understanding of the artefacts and compare him/herself to others. This goal of making awareness explicit is similar to the awareness that has been extensively studied in the human-computer interaction community. One notable issue is that it would be preferable to track annotation traces of documents at the group rather than individual level.

2.2 Information Visualization

With KnowledgeSea II [19], users can explicitly rate each part of a resource by clicking ‘thumbs-up’ (interpreted as a positive rating), ‘thumbs-down’ (interpreted as a negative rating) or ‘question-mark’ (interpreted as a neutral rating). Post-questionnaire evaluations are then conducted to assess users’ perceived subjective perspectives of the recommender. Reffay and Chanier [2] have implemented a Distance Learning Management System which
can identify clusters and cliques among the interactions. They focus on the activity of the class as a whole, not on the individual student.

Counting the number of messages posted by a student is one of the most critical and reliable ways to measure student participation in an online environment. Group activities should be encouraged and emphasized in these e-learning environments [3].

3. Current Stage of Our Work

At this stage, we have finished the algorithmic design of the recommendation as well as carried out an experimental study of pedagogical preferences of actual students in recommending and choosing papers [6, 7, 8]. Our next step is to incorporate the recommendation algorithm in an overall system design, with a focus on maximizing the social affordance of the paper as outlined in this working paper. These features of a paper should be visualized clearly so as to make the information space affordable to learners as a social environment to navigate and socialize.

References