Laying a Foundation for the Graphical Course Map

Linda DuHadway
School of Computing
University of Utah
Salt Lake City, UT 84112 USA
linda.duhadway@usu.edu

Thomas C. Henderson
School of Computing
University of Utah
Salt Lake City, UT 84112 USA
tch@cs.utah.edu

ABSTRACT

The learning managements systems (LMS)s that are widely used to provide access to educational opportunities on the Web are limited by a text based, linear presentation of course materials and the standard temporal restrictions in the traditional classroom. Making a fundamental change in how course materials are presented and interfaced with can make educational opportunities available to a broader spectrum of people with diverse abilities and various circumstances. We have developed a graph-based approach to presenting the learning materials of a course using a system called **ENABLE** [6, 7] with three major goals: (1) facilitate restructuring a set of synchronous classroom materials into a dynamic online system, (2) provide algorithms to analyze and enhance student performance as well as provide insights to the instructor concerning the efficacy of the learning items and their organization, and (3) identify ways to use data from an existing linear, temporal based course presentation to train predictive models for a course that allows individual flexibility in the ordering of the material. This work demonstrates the possibility of presenting course materials in a graphical way that expresses important relations and provides support for manipulating the order and timing of those materials. The graphical course map adds a new approach to making education accessible to people from many different spectrums of ability that respond and interface better with visual representations and those who will benefit from the removal of temporal limitations.

Keywords

Accessibility, Artificial Student Agents, Graphical Course Mapping, Interactive Learning Environments, Personalized Learning, Student Model Calibration, Web-based Learning

1. INTRODUCTION

The availability and accessibility of education over the Web has increased but barriers remain [1, 8, 11, 13]. Current learning management systems (LMS)s display learning materials in a textual, linear format primarily based on chronology. This presentation of material provides a limited view of a course and imposes unnecessary restrictions to online education. Expanding the delivery of learning material to include a graphical course map can increase the information that is available and make that information accessible to a larger number of diverse consumers – especially improving the experience for those who respond to visual representations. Removing the dependence on temporal ordering provides new possibilities for students that may not succeed in the traditional approach of having all students move through learning items at the same time and at the same rate. Making the organization and presentation of educational information visual and flexible makes it available to a broader spectrum of users improving their opportunities for inclusion and success.

Our research focuses on the possibilities of presenting learning materials in a graphical course map. This has led to many discoveries about the opportunities for enhancing the information available to students and educators. The development of a variety of course maps has identified new ways to organize and present learning materials and restructure their delivery to exploit the flexibility of the online setting. The **ENABLE** system provides interactive tools that allow an instructor to focus on meaningful relations between learning items and manipulate a variety of graphical course maps that maintain those relations while introducing new organizational possibilities.

W4A'16, April 11 - 13, 2016, Montreal, Canada

ISBN 978-1-4503-4138-7/16/04…$15.00
DOI: http://dx.doi.org/10.1145/2899475.2899486

2. DISCOVERING EXISTING DATA

The initial stage of this work was to identify what data is currently available and how that data might be used to provide meaningful information for building course maps. Several existing courses were analyzed.

Courses may contain many different instructional materials.For our purposes, these various materials are referred to as *learning items* and a group of learning items is a unit. Each learning item has its own characteristics such as title, due date, content, delivery method, and whether or not it is graded. The sample courses use Canvas, a commercially available LMS. Canvas has a well documented application program interface (API). Using this API allows programmatic access to the data available about the learning items.

Using this data and text analysis methods **ENABLE** is able to identify some relations between learning items. The primary relation available in the LMS is the temporal *precedes* relation. This relation expresses that one learning item comes before another learning item. Another relation commonly found in the LMS is the *includes* relation. The *includes* relation expresses that a unit *includes* a specific learning item and provides information about the existing organization of the course. When the Module tool in Canvas is used to organize a course this relation can be identified from the data. Otherwise additional information from the instructor is needed to inform this relation.

Engaging the instructor provides another source of information that allows the system to identify additional relations such as *occurs in* and *prerequisite* relations. The *occurs in* relation expresses that a specific topic *occurs in* a learning item. The *prerequisite* relation expresses that there is educational value in doing one learning item before another.

3. BUILDING THE COURSE MAP

The information about the learning items and the relations is used to create the course map. A *course map* is a graph, $\mathcal{M} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the set of learning items and topic nodes, and $\mathcal{E}$ is the set of *precedes*, *topically precedes*, *prerequisite*, *occurs in* and *includes* edges (relations). Then the *class map* is $\mathcal{C} = (\mathcal{L}, \mathcal{R})$, where $\mathcal{L} \subset \mathcal{N}$ is the set of learning item nodes and $\mathcal{R} \subset \mathcal{E}$ is the set of edges. A *path* of length k is any legal sequence $P = \{n_1, n_2, \ldots, n_{k+1}\}$, where $n_i \in \mathcal{L}$ and $\neg\exists i, j \ni n_j$ *prerequisite* n_i and $i < j$. Let P^S be the set of nodes in the path P.

Notice that the limiting relation in the path is only the *prerequisite* relation. This provides a wide variety of paths to interact with the learning items. This kind of flexibility does not usually exist in the traditional class setting. In an online educational setting the temporal limitations of the *precedes* relations need not be enforced. Making this shift to allow varied paths through the course material changes how both the educator and the student view a course. Tools described here provide the mechanisms to support such variation.

Once the temporal restraints of the course have been removed, the opportunities for restructuring have increased. The **ENABLE** system can display a graph based on the *occurs in*, *includes*, and *prerequisite* relations without the chronological restraints.

The course map display is designed in such a way that the nodes can be moved about. As a node is moved any connecting edges move with it. Keeping these connections intact during moving preserves the integrity of the graph structure and maintains the relations between learning items. This manual manipulation of the course map provides a way to see the course with many different layouts. The learning items can be organized by topic, exam, learning item type, prerequisite chains, etc. This provides the instructor, and potentially students, the opportunity to explore and discover possible paths through the course material.

4. FACILITATING CHANGE

This work has shown that the data available in the LMS can be used to generate a graphical representation of a course. By gathering additional information from an instructor or other course expert the graphical representation can be expanded to provide additional information and more meaningful relations. This graphical course map provides a new way to see the course materials and how they are related to each other which provides valuable information. The graphical course map is more powerful than the information it can contain. It has the potential to be a mechanism for fundamental change in how education is delivered.

Traditionally courses have run over a specific time frame and are delivered in the same order and timing to all students regardless of ability or circumstance. By adjusting the attachment to a linear, temporally based approach the reach of education to people outside the traditional classroom can be expanded. Currently some see the online course as a way to include people with severe disability in the world of education [10], but many online courses unnecessarily bring the limitations of time and order with them. To decrease this limitation and expand the educational opportunities provided on the Web, a fundamental change needs to occur. Educators and students will need to view the linear, time oriented presentation of a course as an unnecessary limitation and expand their thinking to include alternative approaches.

The graphical course map is a mechanism to support this change in perspective. Simply presenting the learning items in a graphical way allows the educator to see the course differently. Secondly, restraints can be reduced by removing unnecessary connections and focusing on relations that are beneficial to the educational process. The course map can then be manipulated to illustrate new ways to organize the material while keeping these meaningful connections in place.

5. TESTING THE POSSIBILITIES

Allowing students to move through the learning items in various orders introduces an entirely different component to a course. To explore the theoretic impact of such a change artificial student agents, probability models, and calibration techniques were implemented.

5.1 Artificial Student Agents

Detailed student models were developed in order to analyze the relation between the learning item organization and student performance. These artificial student agents can traverse the course map in a variety of node sequences. Different limitations were placed on individual student agents based on their own characteristics of intelligence, work ethic, background, and distractibility. However, the only limitation imposed by the course map is prerequisite relations. A learning item $l \in \mathcal{L}$ is accessible once all the prerequisite learning items have been visited. The agent determines which item to attempt and how well they do on each learning item they visit including the option to apply no effort and receive a zero score. How the agent performs is based on the actual score data for the course that is being analyzed.

A trace of the order the learning items are visited is recorded as each agent moves through the learning items. These learning agents demonstrate a large variety in the order in which the learning items can be attempted. (For more information about these learning agents refer to the author's previous work in [7]).

For use with the estimation method described in Section 5.2, learning agents were created that implement the concept of mastery. A learning item, $l \in \mathcal{L}$, also has an associated difficulty level in the range [0,100] in our experiments. At each time step the agent specifies how much time of the total allotted is to be spent on each accessible learning item; this constitutes an *action*. Agents may implement different learning tactics and their respective learning performance traces may then be compared.

To demonstrate the performance of an artificial student

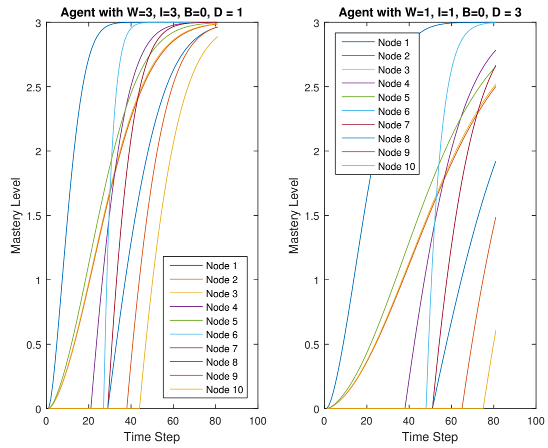

Figure 1: Learning Curves for Agents with Different Abilities.

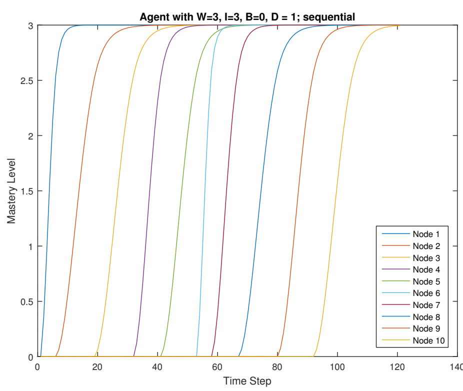

Figure 2: Learning Curves for Agent with Two Different Learning Tactics.

agent consider a class map, $\mathcal{C}$, with 10 learning items. Suppose the agent has characteristics $W = 3$, $I = 3$, $B = 0$, and $D = 1$. Then, a learning curve plot for $Agent_1$ on $\mathcal{C}$ is shown in Figure 1 (left), while an $Agent_2$ with $W = 1$, $I = 1$, $B = 0$ and $D = 3$ is shown on the right side of the Figure 1 (right). $Agent_1$ has achieved almost perfect mastery of all ten learning items by step 80, whereas $Agent_2$ has only mastered a few items in the same time.

Note that learning curves are also a function of the learning tactics of the agent. Suppose that $Agent_1$ modifies its approach to focus on individual items until they are mastered before moving on to the next available item. The resulting learning curve is shown in Figure 2 which illustrates that items are mastered sequentially and it takes longer to learn all ten items than the equal time strategy. [Note that this may also suggest that a linear organization of the course material slows learning!]

5.2 A Learning Model

In [7], we defined the notion of *mastery* of a learning item as a random variable ranging from 0 to 1. This demonstrated the use of a linear learning model combined with a Kalman Filter to obtain an optimal estimate of student mastery of the learning items in a course graph based on combining the model prediction with a measurement (i.e., a grade) correction.

In the present work we propose a more refined learning model which includes a parameter, the *learning coefficient*, and compare three ways to estimate it: (1) direct inverse, (2) iterative least squares (as introduced in [5] and used in [12], and (3) the Extended Kalman Filter (see [14] for a detailed introduction to Kalman Filter methods). This is called either *model parameter calibration* or *parameter estimation*.

The estimation method is based on the use of a class graph which describes the organization of the learning material, a set of artificial student agents with an associated learning model, and a mechanism for the class graph traversal. A wide variety of user models have been proposed for interactive learning environments; e.g., see [2, 3, 4, 9]. We have opted to use a more basic and general model of learning as described in [15]:

$$x_i^{t+1} = M - (M - x_i^t)e^{-k_i s_i^t} + \epsilon \tag{1}$$

where x_i^t is the mastery level of learning item i at time t, M is the maximal mastery level (which we set to 3 in our experiments), k_i is the learning coefficient for the student on learning item i, s_i^t is the cumulative time spent on learning item i, and $\epsilon \sim \mathcal{N}(0, \sigma^2)$. σ^2 is the variance in the learning model process. The learning coefficient k_i is a function of the agent and the learning item:

$$k_i = \frac{\frac{W+I+B}{D}}{\alpha_i} \tag{2}$$

where α_i is the difficulty of learning item i.

The learning coefficient provides a way to compute the relative difficulty of each learning item. This can provide insight to help the instructor better balance the student work load.

5.3 Probability Models

A probability model provides a way to make predictions. Predictions can be used to inform students and educators about possible outcomes. Models can be generated with the data available in the existing course. Can data from a linear, temporal based course be used to predict outcomes for a course that allows different paths through the learning material?

To answer this question several probability models that predict grades on learning items were created. The models are trained using the existing score data. Many of the models are able to predict individual scores with over 70% accuracy. They can also be sampled to produce data that has a distribution similar to the original data. These models can be restricted to only *prerequisite* relations in the existing data and still produce results with similar accuracy. Since the *prerequisite* relations are enforced when traversing the course map this demonstrates that existing data from a linear, temporal based course can be used to predict outcomes for a course that allows more variation.

Table 1 shows a varying degree of accuracy in making score predictions with three of the probability models. The *precedes one* considers the score of the first preceding learning item, the *precedes three* considers the scores of the first three preceding items, and the *prerequisites* considers all the prerequisites for the specified learning item. Restricting the parents in the Bayesian network and the features in the linear and mixed linear methods to prerequisites only reduces the accuracy of the predictions by 2%-5%.

Table 1: Comparing Grade Accuracy Between Models.

Model Type	Dependencies	Grade Accuracy
Mixed linear	Precedes One	75%
Mixed linear	Precedes Three	77%
Mixed linear	Prerequisites	72%
Bayes Net	Precedes One	75%
Bayes Net	Precedes Three	73%
Bayes Net	Prerequisites	73%
Linear	Precedes One	72%
Linear	Precedes Three	72%
Linear	Prerequisites	67%

6. CONCLUSIONS

This work lays a foundation for the creation of general purpose graphical course mapping tools. It demonstrates the possibility of generating such a map using currently available data. The manipulatable course map can support educators as they transform a synchronous, temporal based course presentation to one without the same limiting temporal restraints. It allows the possibility of individual students moving through the course materials in a variety of orders and time frames.

Several automated student agents have been developed and used with learning models to consider how basing the ordering of learning materials on *prerequisite* relations might impact the learning process. This initial investigation found that there were many ways to navigate through the learning material of several sample courses while enforcing the *prerequisite* relation.

Predictive models were produced and used to demonstrate that data from existing linear, temporal based courses could be used to train predictive models. These predictive models could be limited to *prerequisite* relations and still produce accuracy results that are just slightly less than when *precedes* restraints are included in the data. This provides a way to generate recommendation systems for students and educators using a more flexible delivery method.

Phase I of this work provides a solid foundation for the creation of graphical course mapping systems. For such a system to become widely useful, an interface is needed that incorporates the relations discovered and the recommendations available through the predictive models. Phase II is the next step of this work and includes (1) creating a rich graphical user interface that improves both the quality and quantity of student and teacher interaction with the learning material, and (2) conducting user testing at all stages of the system design, development, and testing to identify the usability and accessibility of the interface and make revisions based on the results. Such an interface could then be embedded in the LMS for student and faculty use on the Web or in mobile devices.

7. REFERENCES

[1] I. E. Allen and J. Seaman. Grade Change: Tracking Online Education in the United States, 2013. *Babson Survey Research Group and Quahog Research Group, LLC. Retrieved on*, 3(5):2014, 2014.

[2] P. Brusilovsky, S. Sosnovsky, and O. Shcherbinina. User Modeling in a Distributed E-Learning Architecture. In *User Modeling 2005. Lecture Notes in Artificial Intelligence, (Proceedings of 10th International User Modeling Conference)*, pages 24–29. Springer Verlag, 2005.

[3] C. Carmona, G. Castillo, and E. Mil'an. Designing a Dynamic Bayesian Network for Modeling Students' Learning Styles. In *Proceedings of the International Conference on Advanced Learning Technologies*, pages 346–350, Santander, Cantabria, Spain, July 2008. IEEE.

[4] C. Carmona and R. Conejo. A Learner Model in a Distributed Environment. In *Proceedings of the Third International Conference on Adaptive Hypermedia and Adaptive Web-based systems*, Eindhoven, The Netherlands, August 2004.

[5] S. Chapra and R. Canale. *Numerical Methods for Engineers*. McGraw Hill, Boston, MA, 2002.

[6] L. DuHadway and T. Henderson. Informing Change: Course Content Analysis and Organization. In *Proceedings of the Frontiers of Education Conference*, El Paso, TX, October 2015. IEEE.

[7] L. DuHadway and T. Henderson. Artificial Student Agents and Course Mastery Tracking. In *Proceedings International Conference on Agents and Artificial Intelligence*, Rome, Italy, February 2016.

[8] C. S. Fichten, V. Ferraro, J. V. Asuncion, C. Chwojka, M. Barile, M. N. Nguyen, R. Klomp, and J. Wolforth. Disabilities and e-Learning Problems and Solutions: An Exploratory Study. *Journal of Educational Technology & Society*, 12(4):241–256, 2009.

[9] N. Henze and W. Nejdl. Student Modeling in an Active Learning Environment Using Bayesian Networks. In *In Proceedings of the Seventh International Conference on User Modeling, UM99.* Springer, 1999.

[10] S.-M. Lee. Universal Design in Online and Higher Education Through ICT. In *Proceedings of the 11th Web for All Conference*, W4A '14, pages 9:1–9:1, New York, NY, USA, 2014. ACM.

[11] T. Müller. Persistence of Women in Online Degree-Completion Programs. *The International Review of Research in Open and Distributed Learning*, 9(2), 2008.

[12] M. Myers, A. Jorge, and D. G. Walker. A Comparison of Extended Kalman Filter Approaches using Non-Linear Temperature and Ultrasound Time-of-Flight Measurement Models for Heating Source Localization of a Transient Heat Transfer Problem. In *Inverse Problems, Design and Optimization Symposium*, Joao Pessoa, Brazil, August 2010.

[13] D. Rover, Y. Astatke, S. Bakshi, and F. Vahid. An Online Revolution in Learning and Teaching. In *Proceedings of the Frontiers of Education Conference*, Oklahoma City, OK, October 2013. IEEE.

[14] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, Cambridge, MA, 2005.

[15] W. Zangwill and P. Kantor. Toward a Theory of Continuous Improvement and the Learning Curve. *Management Science*, 44(7):910–920, July 1998.