Refining Learning Maps with Data Fitting Techniques: Searching for Better Fitting Learning Maps

Seth Adjei, Douglas Selent, Neil Heffernan
Worcester Polytechnic Institute
{saadjei, dselent, nth}@wpi.edu

Zach Pardos
UC-Berkeley
pardos@berkeley.edu

Angela Broaddus, Neal Kingston
University of Kansas
{broaddus,nkingsto}@ku.edu

ABSTRACT
Learning sciences needs quantitative methods for comparing alternative theories of what students are learning. This study investigated the accuracy of a learning map and its utility to predict student responses. Our data included a learning map detailing a hierarchical prerequisite skill graph and student responses to questions developed specifically to assess the concepts and skills represented in the map. Each question aligned to one skill in the map, and each skill had one or more prerequisite skills. Our research goal was to seek improvements to the knowledge representation in the map using an iterative process. We applied a greedy iterative search algorithm to simplify the learning map by merging nodes together. Each successive merge resulted in a model with one skill less than the previous model. We share the results of the revised model, its reliability and reproducibility, and discuss the face validity of the most significant merges.

Keywords
Learning Maps, Iterative Search, Cognitive Modeling, Skill Graph

1. INTRODUCTION
Cognitive models are used to represent how one's knowledge may be organized. As such, they contain descriptions of component pieces of knowledge and connections among the components to indicate how understanding develops in a specified domain [4]. Different authors have described various cognitive models, including learning maps [5], learning trajectories [2], and learning hierarchies [3]. Learning maps use linear sequences of learning goals and are useful for instructional planning [5]. A learning trajectory includes a learning goal, a developmental progression defining the levels of thinking students pass through as they work toward the defined goal, and a set of learning activities or experiences that assist students in reaching the defined goal [2]. Learning hierarchies model prerequisite knowledge components in hierarchies, allowing multiple pathways to extend from one prerequisite skill to multiple learning goals [3].

In the present study we examine a small section of the learning map and investigate the effects of permuting the topology of the hierarchy. Skills and concepts are represented by latent nodes in the learning map. Directed edges represent the prerequisite relationship among latent nodes and also represent the relationship between those nodes and their associated test items. We present a simple method for improving the predictive power of the learning map by combining latent nodes.

This work connects with literature on searching for better fitting cognitive models. Several non-hierarchical cognitive models have been developed to represent the relationship between knowledge components (KCs) in the form of prerequisite skill maps. These cognitive models have been developed to help intelligent tutors, as well as experts, determine student mastery of KCs. A number of technical approaches have been developed to evaluate cognitive models developed by domain experts. One approach is Learning Factors Analysis (LFA), developed by Cen, Koedinger and Junker [1] to help the Educational Data Mining (EDM) community evaluate different cognitive models. LFA incorporates a statistical model, item difficulty and a combinatorial search to select the model. Our work is different from the flat Item Response Theory (IRT) models presented in [9] in that IRT does not deal in any way with hierarchical relationships between knowledge components.

In this work we follow the process described by Cen, Koedinger and Junker [1]. This technique can be used to analyze hypothesized learning maps and consider whether small improvements to the model result in a better fit to the data. In this method two different approaches were studied to determine the best skill map from an initial graph. Cen, Koedinger, and Junker suggested three types of operations, i.e., merges, splits, and adds [1]. However, in this study, we used only merge operations given the already highly granular quality of our initial, subject matter expert derived learning map.

2. Initial Learning Map
This study examined a section of the learning map containing 15 concepts and skills related to understanding integers. The map was developed using mathematics educational literature describing how students learn to understand and operate with integers. The set of integers includes the whole numbers and their opposites, presenting many students their first exposure to negative numbers [6].

The data for this study was gathered from responses of 2,846 students to 25 test items aligned to 15 skills. All of the test items were multiple choice questions, with four answer options per question. Each skill was assessed by one or more items. As part of the test development process, subject matter experts confirmed the alignment of each item to its associated skill, meaning that the item was judged by experts to evoke the intended skill. Furthermore, due to the hierarchical structure of the learning map, items associated with skills lower in the learning map were assumed to be more difficult, i.e., require more skills, than items associated with skills higher in the learning map.

3. Experiment

In all of the experiments our sole manipulation of the map was to merge latent nodes. A merge operation occurred when two skills adjacent to each other in the map were combined into one skill. Items from both skills that were merged were reattached to the new single skill. The prerequisites of the constituent skills became prerequisites of the merged skill and the same applied to the post-requisites.

To evaluate the models, we used per student per item cross validation with 5 student folds and 3 item folds. Our student and item folds were chosen randomly for the evaluation; however each item fold consisted of the same random partition of items. More details about how the cross-validation was done as well as other details on the algorithms used in this experiment can be found in the technical document [8]. We used the Root Mean Squared Error (RMSE) of the predictions to evaluate the results of the experiment. A lower RMSE means the model is performing with a higher accuracy.

Figure 1 shows a graph of the results from the iterative search. The search started at iteration 1, which was the initial skill map consisting of 15 skills before any merges were applied to it. The search ended at iteration 15, which is a graph consisting of just one skill with all the items attached to that one skill. The best models from each iteration are shown in Figure 1. We used RMSE to choose the best model at each iteration and to guide our search. As the number of iterations increases, the number of skills decreases since skills are merged.

Performance of the Number of Skills

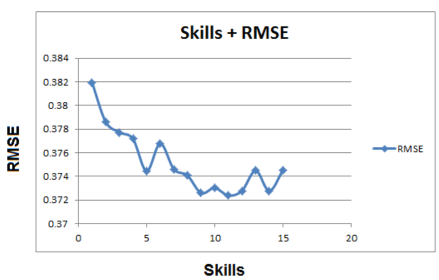

Figure 1: Performance for each number of skills. Each merge operation reduces the number of skill by 1. After iteration 1, there are a total of 15 skills. After iteration 15, there is only one skill

The results show that the best RMSE obtained was from the 11-skill map at iteration 4 with an RMSE of 0.372. This is slightly better than the original skill map with RMSE of 0.375. The 11-skill map has a small but significant improvement ($p < 0.01$) from the original skill map. The effect size was negligible (0.01).

4. Contributions, Conclusions and Future Work

The main contribution of this paper is the provision of a greedy algorithm that simplifies learning maps. We showed that this simplification is possible without losing the predictive powers of the learning map. Even though this simplification could be done by hand, this algorithm will be useful in situations where the learning map being simplified is large.

This paper presents an initial experiment in this novel area of EDM. Instead of focusing all our attention on the flat IRT model, the community needs to pay a closer attention to and explore models that deal with hierarchical relationships between knowledge components. These studies and contributions thereof can assist domain experts to produce better fitting models which should impact student learning positively.

Since merging skills increased accuracy, these results suggest that the original skill map was too fine-grained (given the number of questions per skill and the number of students who took the test.). In some cases the test items did not adequately distinguish between the skills that were merged; hence such skills were merged. The results of algorithms like this can help the content experts who are creating skill maps and test items to either reconsider thinking of two skills as separate, or prompt them to write different test items to better distinguish between students that have mastered one of the skills but not the other skill. As future work, we intend to examine the other operations of the LFA method for refining learning maps. These include *splits* and *adds*, which were described earlier.

5. REFERENCES

[1] Cen, H., Koedinger, K. R., & Junker, B. (2006). Learning Factors Analysis: A general method for cognitive model evaluation and improvement. In M. Ikeda, K. D. Ashley, T.-W. Chan (Eds.) Proceedings of the 8th International Conference on Intelligent Tutoring Systems, 164-175. Berlin: Springer-Verlag.

[2] Clements, Douglas, and Julie Sarama. (2004). "Learning Trajectories in Mathematics Education." Mathematical Thinking and Learning An International Journal 6:81-89.

[3] Gagné, Robert. 1968. "Learning Hierarchies." Educational Psychologist 6:1-9.

[4] Gierl, Mark, Changjiang Wang, and Jiawen Zhou. 2008. "Using the Attribute Hierarchy Method to Make Diagnostic Inferences about Examinees' Cognitive Skills in Algebra on the SAT." The Journal of Technology, Learning, and Assessment, 6(6):4-49.

[5] Popham, W. James. 2011. Transformative Assessment in Action: An Inside Look at Applying the Process. Alexandria,VA: ASCD.

[6] Van de Walle, J. A., Bay-Williams, J. M., Karp, K. S., Lovin, L. H. (2014). Teaching Student-Centered Mathematics: Developmentally Appropriate Instruction for Grades 6-8 (2nd ed.). Upper Saddle River, NJ: Pearson Inc.

[7] van der Linden, W. and Hambleton, R. (Eds.). Handbook of Modern Item Response Theory. New York, NY: Springer-Verlag (1997).

[8] Learning Maps Dataset and Evaluation Code https://sites.google.com/site/assistmentsdata/kansas-project

Last accessed on February 25, 2014