Predicting Student Performance on Post-requisite Skills Using Prerequisite Skill Data: An Alternative Method for Refining Prerequisite Skill Structures

Seth A. Adjei
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609-2280
saadjei@wpi.edu

Anthony F. Botelho
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609-2280
abotelho@wpi.edu

Neil T. Heffernan
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609-2280
nth@wpi.edu

ABSTRACT
Prerequisite skill structures have been closely studied in past years leading to many data-intensive methods aimed at refining such structures. While many of these proposed methods have yielded success, defining and refining hierarchies of skill relationships are often difficult tasks. The relationship between skills in a graph could either be causal, therefore, a prerequisite relationship (skill A must be learned before skill B). The relationship may be non-causal, in which case the ordering of skills does not matter and may indicate that both skills are prerequisites of another skill. In this study, we propose a simple, effective method of determining the strength of pre-to-post-requisite skill relationships. We then compare our results with a teacher-level survey about the strength of the relationships of the observed skills and find that the survey results largely confirm our findings in the data-driven approach.

Categories and Subject Descriptors
K.3 [Computers and Education]: Miscellaneous

General Terms
Theory, Experimentation

Keywords
Prerequisite Structures, learning maps, skill relationships, refinements, PLACEments

1. INTRODUCTION
Prerequisite skill structures represent the ordering of skills in a given knowledge domain. The learning sequences represented in the prerequisite skill structures have become an area of interest over the past few years. As a prelude to the objective of learning prerequisite skill structures from data, Tatsuoka [12] developed and proposed the Q-Matrix, a structure that represents the mapping of items on a test to specific skills. Others have built on this structure to find relationships between the skills and items represented in the Q-matrices [3,11] or proposed methods for refining Q-Matrices [7]. Brunskel presented preliminary work in which she used students' noisy data to infer prerequisite structures [4]. Additionally, Scheines, et al. [11] present an extension of a causal structure discovery algorithm in which the assumption of pure items are relaxed to reflect real data, and use that relaxed assumption to infer prerequisite skill graphs from students' response data.

The focus of other researchers in the community has been on refining the prerequisite structures developed either by domain experts or through data mining approaches, as used by Barnes [3]. Cen, et al. [5] proposed Learning Factors Analysis (LFA) as a method for refining cognitive models. Their approach includes statistical techniques, human expertise, and combinatorial search to refine cognitive models. Following the proposals made by Cen et al. in [5], Adjei et al. [1] developed a combinatorial search algorithm based on LFA and found simplified prerequisite structures, which have equally good predictive power as the originals.

Desmarais, et al. [6] introduced a method for determining partially ordered knowledge structures (POKS) from student data. The main idea behind this approach is to compare pairs of items in a test in order to determine any interactions existing between each pair. The interactions serve as a basis for determining the relationship between the skills represented by the items. Pavlik and his colleagues applied POKS to analyze item-type covariances and proposed a hierarchical agglomerative clustering method to refine the tagging of items to skills [9] and later proposed Learning Factors Transfer Analysis [10] as a means for generating domain models. Adjei and Heffernan [2] used randomized control experiments to identify links within prerequisite skill structures that require further scrutiny. All of this effort that has been expended in the quest to find skill structures from data have yielded varied degrees of success.

The desire to find the best representation of skills (i.e., the prerequisite skill structure) is important for a number of reasons. It informs domain experts about the optimal sequencing of instruction in order to achieve the best tutoring for students. Additionally, this should help researchers in the education research community to better model students' knowledge and performance in intelligent tutoring systems more accurately.

This current study proposes a simple method for identifying problematic links in a prerequisite skill structure, pointing domain experts to the ordering of instructions that may be creating problems for students. In this study we use linear regression of students' performance on items presented to students in the order of a given prerequisite skill structure and make suggestions about the strength of the relationships between the skills.

This paper starts out by describing PLACEments, the adaptive testing system from which data was collected for use in this study. This is followed by a description of the methods we employed and the results of the studies. We then present the results of a teacher survey that we conducted and compare the results of the survey with the findings of our data mining. The paper concludes with a discussion of the results and possible future work in this area.

1.1 Brief Introduction to PLACEments

PLACEments, a free mathematics adaptive testing system, is a feature of ASSISTments (a free web-based Intelligent Tutoring System (ITS)). When assigning a PLACEments test, an initial set of skills are selected for the test. Students are tested on the initial set of skills and depending on their performance, the system traverses a skill graph to present problems from the prerequisite skills of the initial set of skills. The test adapts to the student's performance as well as the underlying prerequisite skill graph. If a student performs poorly on an item in the test, they are presented with items from the prerequisite skills required to solve the original problem. PLACEments uses a prerequisite skill structure created by one of the experts who developed the Common Core Standard for mathematics. Portions of this structure are currently being used by websites like AchieveTheCor.org [http://www.achievethecore.org/coherence-map/]. The developers of the site call it the Coherence Map. (A small portion of this structure is shown in Figure 1.)

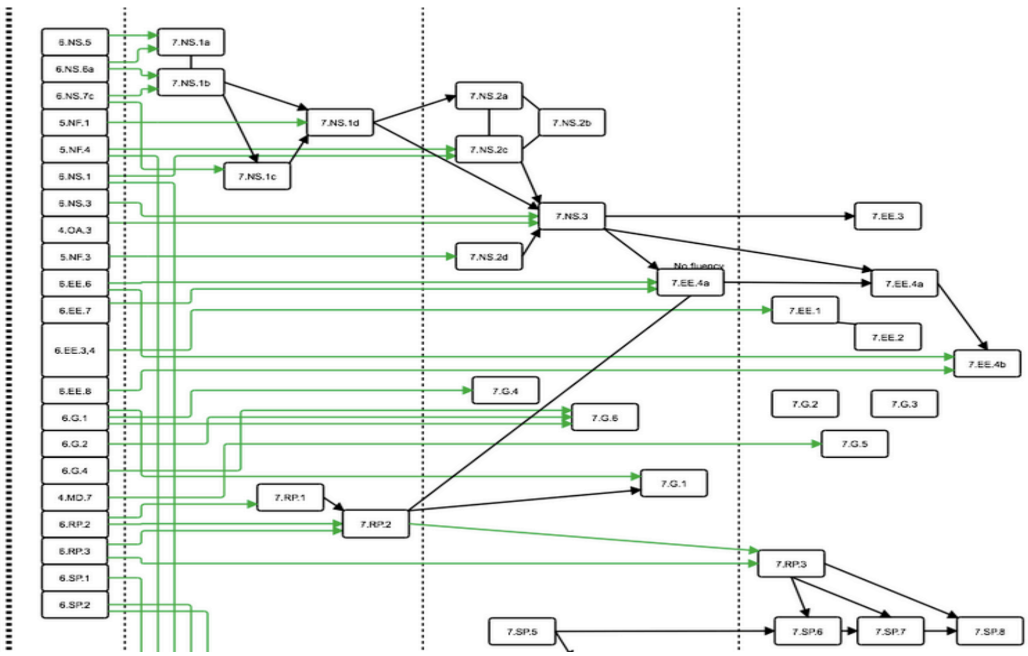

Figure 1. A prerequisite Skill Graph. Each rectangular box in the structure represents a Common Core Standard for Mathematics. The green arrows in the graph represent links to prerequisite skills from lower level grades while the black link grade level prerequisite skills. Arrows point to the post-requisite skills. A link between nodes that does not have an arrowhead represents a relationship between the two skills, but not a pre- or post-requisite relationship. Thus, they do not specify an order of instruction.

PLACEments has an additional feature that assigns remediation assignments to students who perform poorly on a test. These remediation assignments are intended to build the students' understanding of the skills they performed poorly on, during the test. The remediation assignments are released in the order of the arrangement of skills in the prerequisite skill structure. Students are assigned lower grade level prerequisite skills first, and until they complete those remediation assignments, post-requisite skills-related remediation assignments are not released. This ensures that the students gradually build on their knowledge of skills until they eventually reach a desired level of mastery of the skills in the given domain.

To illustrate how PLACEments works, Figure 2 shows a hypothetical prerequisite skill graph where the letters A through H each represents a skill. The graph additionally shows a typical configuration of a student's navigation through the prerequisite skill structure in the process of taking a PLACEments test. Skills A, B and C are the initial skills assigned on the test.

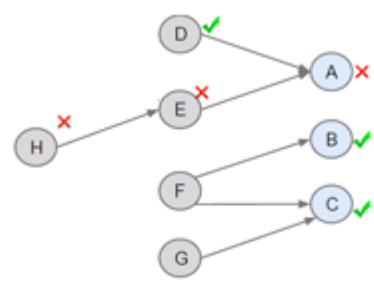

Fig 2. A Typical Student's navigation in a PLACEments test

In this case the student performed poorly on skill A and so is asked questions for skills D and E. Since the student could not demonstrate understanding on skill E, he is further asked questions from skill H, which he performs poorly on as well. PLACEments creates remediation assignments for each of the skills the student performs poorly on (A, E and H). For this particular example, the remediation assignment for skill H is released before any other remediation assignments are released. The assignment for skill E is released after the student completes that previous skill's assignment.

For the purpose of this study, we focus only on the remediation assignment management feature. This is the feature that provides us with data for determining how strong prerequisite skill relationships are. The remediation assignments are typically assignments in which students practice a number of similarly designed problems to help them master a particular common core skill. In the course of the assignments, students are allowed to ask for help (in the form of hints) as they progressively answer the questions. The student is deemed to have mastered the skill if he/she correctly answers *n* consecutive problems in the assignment without asking for hints. The value of *n* typically ranges from three to five depending on the designer of the problem set. If after a set number of problems (typically called the daily limit), the student is unable to reach the mastery criterion, the system pauses the practice session until the next day when the student can continue with the assignment.

2. Methodology

2.1 Dataset

The remediation assignment feature of PLACEments served as the source of data for the current study. The dataset includes students' performance on remediation assignments. There were 495 prerequisite skill links from the prerequisite skill structure described above (fig 1). In this study we focused our attention only on skills that have exactly one prerequisite skill. Of the 104 skills that have exactly one prerequisite skill, we had 14 of the links that had data for a minimum of 50 students. For each of the prerequisite skill links examined, there was an average of 120 students who were assigned remediation assignments of both the prerequisite and post-requisite skills of the link.

Each row in the dataset has a student's performance on the pre- and post-requisite skills (measured by the percent correct of the Skill Builder, and the number of items it took them to complete the Skill Builder typically referred to as the student's mastery speeds) and the student's prior performance on all problems in ASSISTments. The latter is to help us account for the student's knowledge level. The data set also includes the skill difficulty values for both the pre- and post-requisite skill. These difficulty values are the percent correct for all the items tagged with that skill in ASSISTments. Table 1 shows a sample of the dataset that was used for this study. Each row in the dataset represents a student's' performance on the remediation assignments related to a given PLACEments test. If the student had a similar pair of assignments in another placements test, that information was ignored because we did not want to duplicate the data for a given student. In all, the dataset had 5803 instances of student's performance on pre- and post-requisite skills, involving 1567 students who have completed placements tests.

2.2 Regression Models

We ran linear regression to predict students' performance on the prerequisite skill. A model was built for each of the 14 prerequisite links in the skill graph. In each model, we accounted for the students' overall performance and the difficulty of both the pre- and post-requisite skills, so that our findings will not be influenced by any of those factors. Specifically, our predictors were student's prior performance, mastery speed of the pre-requisite skill, and the difficulty of both the pre- and post-requisite skills. The dependent variable was the mastery speed of post-requisite skills.

Table 1. Sample data set. [1]

SID	PsSk	Pre Sk	Pos MS	Pre MS	StPr	Pre Dif	Pos Dif
23412	57	50	4	5	0.75	0.32	0.40
24321	87	50	3	5	0.86	0.58	0.67

The following equation illustrates the regression model learned from the data for each of the links:

$$ms_{i,j} = \alpha_i + \beta_k * ms_{k,j} + \gamma_{i,k} K_j + \rho_k dif_k + \sigma_i dif_i \; \ldots (1)$$

where $ms_{i,j}$ represents student j's mastery speed for post-requisite skill i

$ms_{k,j}$ represents student j's mastery speed for prerequisite skill k

α_i represents the intercept for the regression estimated from the data

β_k represents the coefficient of mastery speed $ms_{k,j}$

$\gamma_{i,k}$ represents the beta value for the students' prior knowledge K_j

dif_k and dif_i represent the difficulty for skills j and k respectively

ρ_k is the coefficient of the difficulty of the post requisite skill

σ_i is the coefficient of the difficulty of the post requisite skill

We considered a link's model only when the model was found to be statistically significant ($p<0.05$) with R-Square above 0.1. All those models with R-Square values below 0.1 were considered to be suggestive of non-existence of a believable link between the two skills. For the models that met the above criterion, a prerequisite relationship was considered to exist when there is a positive standardized beta coefficient for the prerequisite skills mastery speed (i.e., $\beta_k > 0$) and is significant ($p<0.01$ in many cases and $p<0.05$ in a few).

Since outliers in the dataset could skew the results, we used two data transformation methods to minimize the effects of outliers in the dataset. The first method was to winsorize the mastery speeds in which all mastery speeds above 10 had their values set to 10. Skill Builders in ASSISTments have this feature where a daily limit of 10 is set to prevent students from banging their heads when they are unable to master the skill within 10 opportunities. This is the reason we chose 10 as the cut off number in order to fairly account for student performance. More than 80% of the data

[1] The complete dataset can be found at http://tiny.cc/mslinkstrength. SID is the unique student identifier, PsSk is the post requisite skill id, PreSk has the prerequisite skill id, PrMS and PosMS contain the student's mastery speed of the pre- and post-requisite skill respectively, StPr is the students' prior percent correct (an indication of the student' knowledge level), and PreDif and PosDiff is the difficulty of the pre- and post-requisite skills. The column names have been shortened for lack of space.

we used had mastery speeds below 10 so the impact of this transformation was not very significant. The second data transformation method we used was a log transform of the mastery speeds. We then used each of the transformations to predict the correspondingly transformed mastery speeds and present both results in the results section.

In the case of the transformed data, we replaced the raw mastery speeds in the model with the transformed mastery speeds. We run linear regression models similar to equation (1) above with the mastery speeds, $ms_{i,j}$ and $ms_{k,j}$, respectively replaced with the transformed data, $\widetilde{ms_{i,j}}$ and $ms_{k,j}$. Equation (1) in this case becomes:

$$\widetilde{ms_{i,j}} = \alpha_i + \beta_k * \widetilde{ms_{k,j}} + \gamma_{i,k}K_j + \rho_k dif_k + \sigma_i dif_i \quad \ldots\ldots (2)$$

2.3 Teacher Survey

To verify the results of our findings, we ran a survey of 45 randomly selected domain experts and teachers who use ASSISTments and asked about their perceptions of the strength of the 24 prerequisite skill relationships, including the 14 links we studied in the regression study. A sample survey question is shown in Figure 3. The survey had 26 different questions, the first question introduced the survey and the last was complimentary. There was a survey question for each of the 24 prerequisite skill links. For each prerequisite skill link, we presented a sample problem for each of the post- and pre-requisite skills and asked teachers to rate, on a scale of 1 to 7 (1 not important; 7 extremely important), how important it is for a student to know the prerequisite skill to be able to answer the problem from the post-requisite skill. Even though the questions give the impression that we are trying to figure out how related the skills are, we intentionally did not use the terms pre- and post-requisite skills in order not to confuse the respondents, or to point them in a particular direction. A link is considered to exist if the mean of the responses for that link was approximately 5 and a standard deviation of less 1 or less. We then compare the results of the survey with the findings of the study and report on the comparison.

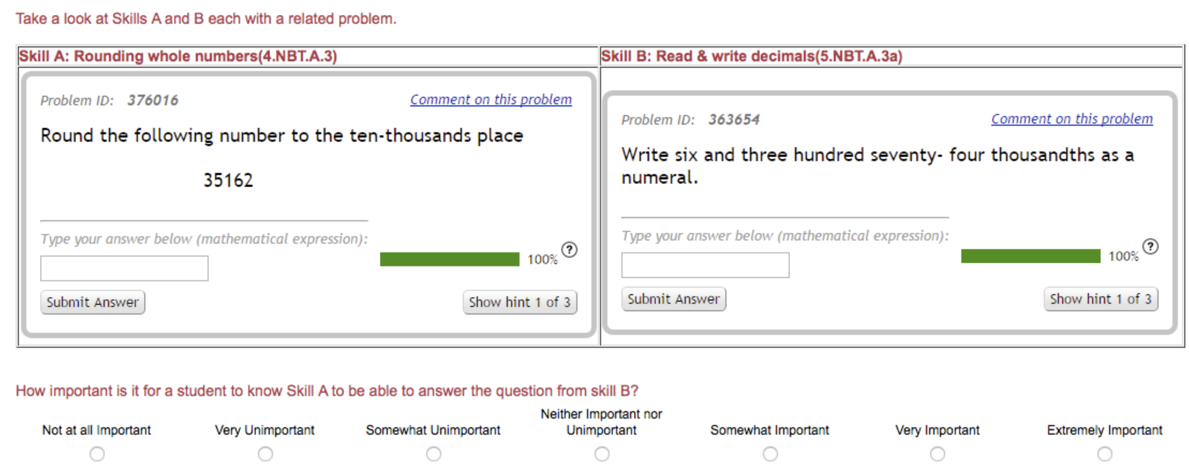

Figure 3. Sample Teacher Survey question

3. Results

3.1 Regression

Table 1 shows some of the links involved in the study and the p-values of the prerequisite skills' mastery speeds predictors in the linear regression model. A link is considered to be valid or strong if the prerequisite mastery speed is determined to be a significant predictor in the regression model and has a positive standardized beta.

Figure 4 shows that several of the links could be found to be problematic and require further scrutiny. When the mastery speeds were transformed to take care of the outliers, the models do a better determination of the good links than was the case when the raw speeds were used. The takeaway from this graph is that we do a better job at finding both good and bad links in a prerequisite structure (and thus refine the structure) when we transform the mastery speeds. By transforming the data, we increase the good links by about 10 percentage points, as shown in Figure 4.

Table 2: A sample of the prerequisite skill links that were considered

Post-requisite Skill	Prerequisite Skill
Adding mixed numbers like denominator	Subtracting fractions like denominator
Subtraction Proper Fractions Unlike Denominators	Addition Mixed Numbers
Addition Proper Fractions Unlike Denominators	Subtraction Proper Fractions Unlike Denominators
Addition Mixed Numbers	Subtraction Mixed Numbers
Volume of rectangular prism without formula	Volume Rectangular Prism
Division Whole Numbers	Divide multi-digit numbers
Subtraction Positive Decimals	Multiplication Positive Decimals
Multiplication Positive Decimals	Division of Positive Decimals

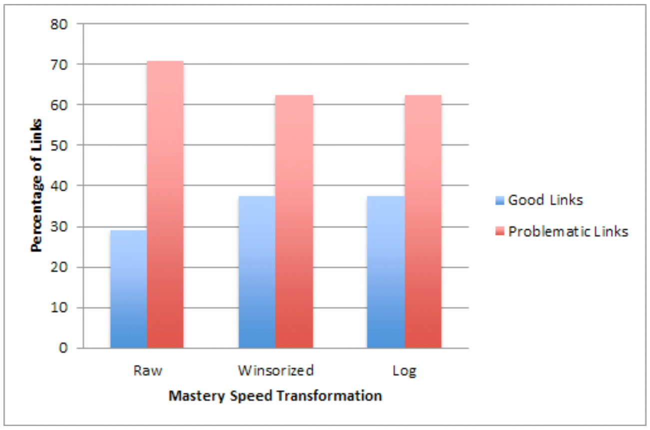

Figure 4. Percentages of identified good and problematic links based on mastery speed transformation methods

The graph in figure 5 shows that of the 24 prerequisite links, the regression method identified 25% of the links (6 links) in which the students' performance on the prerequisite skills significantly predicts their performance on the post-requisite skills irrespective of the data transformation method used. 36% of the prerequisite skills (in 8 links) are significant predictors (p-value < 0.05) of the post-requisite skills if we used any two of the transformation methods. This suggests that a prerequisite skill relationship truly exists between the two skills in each of those links.

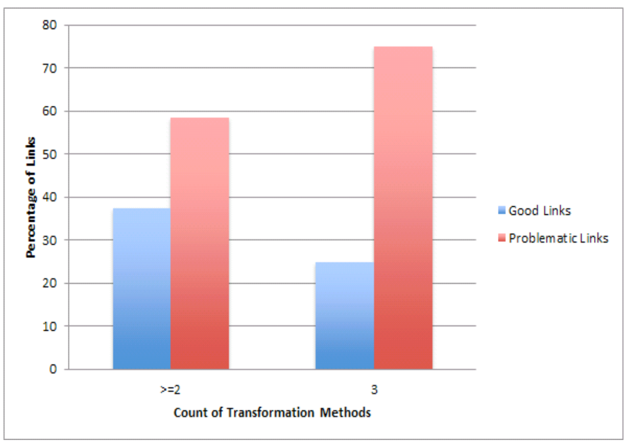

Figure 5. Agreement between mastery speed transformation methods

3.2 Teacher Survey

Of the 45 teachers who were invited to respond to the survey, we received responses from 21 teachers (representing a response rate of 47%). All the respondents completed the survey.

The responses for each of the 24 prerequisite links were averaged per link. Links that have average responses greater or equal to 5 were viewed as important. This is because all possible responses above 4 (i.e., neither important nor unimportant) have a degree of importance ranging from somewhat important (5) to extremely important (7). Therefore if the average response is above 5, we can safely conclude that the prerequisite relationship between the skills in the link exists.

Based on the above criteria, we used the responses from the survey as ground truth and compared our regression models to this ground truth. Figure 6 presents how the regression models compare with the results of the survey. Though this is not a prediction exercise, we like to use some of the metrics for predictions to report on the comparison between the regression models predictions and our survey. We used precision, recall, and accuracy metrics.

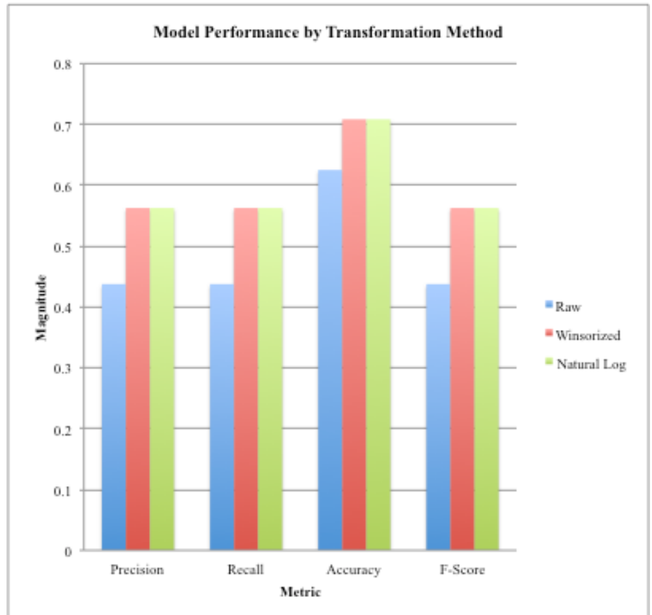

Figure 6. A comparison of model predictions to teacher survey about link strength

An interesting finding from a comparison with the survey is that the accuracy of the regression models did not change despite the type of transformation method used.

Figure 7 presents another perspective of the performance of the method proposed. The graph shows the number of links that were correctly identified from the study grouped by transformation method, given that we accept the results of the survey as ground truth. In general, we see that the method does a good job of identifying a good number of links, though the accuracy numbers from figure 6 could be better.

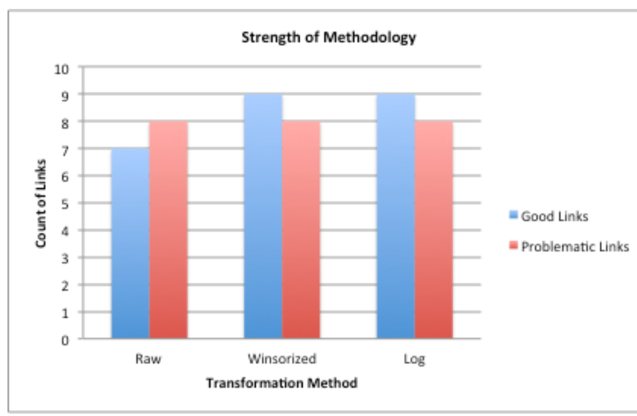

Figure 7. Strength of Methodology: This graph shows counts of the good and bad links correctly identified by the method grouped by data transformation method.

4. Discussion of Results

Prerequisite skill structures in any knowledge domain are very important for instruction and for preparing students for future learning. Almost every knowledge domain has one or a couple, which are created by domain experts. It is important to note that several of these prerequisite skill structures need to be refined.

Data is currently being generated that affords us the opportunity to use data-centered methods to refine these structures.

In this study we used data generated from PLACEments, which is an adaptive testing feature of ASSISTments, to propose one method for refining these structures. In this current study, we used a simple linear regression method in which we use the student performance on the prerequisite skill to predict their performance on the post-requisite skill. We found that for some of the links, this method was effective at identifying both good and bad links in the structure. Comparing the results of the method with the survey provided a ground truth with which to compare the findings from the study. The results have shown that if we have performance data, in the form of mastery speeds, we can achieve more accurate results by transforming the dataset in some format in order to take care of outliers that can easily ruin the findings. [8] The methodology affords prerequisite skill structure creators i.e., domain experts, the opportunity to identify and refine the order of the skills in these structures.

It must however be noted though that the method was not perfect. A few of the links could not be correctly identified. Additionally, the criterion for determining whether a regression model is worth examining is relatively low. In view of these, further studies are required to ascertain the reasons behind that finding and to propose refinements of the method. There could be interaction effects and other relevant predictors that have been ignored, but which may be necessary to ensure that the models' predictions are more accurate.

5. Contribution

The main contribution of this study is the provision of a simple and effective linear regression based method for refining prerequisite skill structures. With this method, we are able to identify problematic arcs in the structure and make these findings available to domain experts who will then use this information to further refine the maps. Additionally, we have introduced a system that provides us with a very good source of data for refining prerequisite skill structures.

6. Conclusion

Several authors in the educational data mining and learning analytics community have attempted to learning prerequisite structures from students response data. Others have looked for methods to refine already existing, domain-expert-made prerequisite skill structures using methods like LFA, etc. It has so far been difficult to get datasets that present students' response data in the order of their underlying students performance. This paper uses such a dataset available from PLACEments, an adaptive testing system that traverses a prerequisite skill structure for item selection. The data from students performance on remediation assignments was used to learn the strength of prerequisite skill relationships existing between skills and to make suggestions regarding these arcs.

We have shown that using simple linear regression and with the right dataset, we can tell how strong the prerequisite skill relationship between two skills are and, based on that make suggestions, regarding which links domain experts may need to investigate and refine.

This study has limitations though. The data used for this analysis has come solely from PLACEments. There are not many such adaptive testing systems that generate the kind of data we used in this study. It will be interesting to study another dataset that is generated in the same or similar format as PLACEments does and to apply this simple linear regression method to make statements about the strength of such prerequisite skill relationships.

We see this study as a preliminary step in our quest to find optimal prerequisite structures using student's response data from placements.

7. Acknowledgements

We thank multiple NSF grants (ACI-1440753, DRL-1252297, DRL-1109483, DRL-1316736, DGE-1535428 & DRL-1031398) , the US Dept. of Ed (IES R305A120125 & R305C100024 and GAANN), and the ONR.

8. References

[1] Adjei, S., Selent, D., Heffernan, N., Pardos, Z., Broaddus, A., & Kingston, N. (2014). Refining Learning Maps with Data Fitting Techniques: Searching for Better Fitting Learning Maps. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 413– 414). Accessed on October 31, 2015, from http://educationaldatamining.org/EDM2014/uploads/procs2014/posters/89_EDM-2014-Poster.pdf

[2] Adjei, S. A., Heffernan, N. T (2015) Improving Learning Maps Using an Adaptive Testing System: PLACEments. In Conati, C., Heffernan, N., Mitrovic, A., Verdejo, M.F. (Eds.) *Proceedings of the 17th International Conference* (pp. 517-520), AIED 2015, Madrid, Spain, June 22-26, 2015.

[3] Barnes T. (2005). The Q-matrix Method: Mining Student Response Data for Knowledge. In: *Proceedings of AAAI 2005 Educational Data Mining Workshop*.

[4] Brunskel, E. (2011). Estimating prerequisite structure from noisy data In Proceedings of *International Conference on Educational Data Mining (EDM)*

[5] Cen, H., Koedinger, K. R., & Junker, B. (2005). Learning Factors Analysis: A general method for cognitive model evaluation and improvement. In M. Ikeda, K. Ashley, & T. Chan (Eds.), *Intelligent Tutoring Systems 8th International Conference* (pp. 164–175). Berlin: Springer.

[6] Desmarais, M., Maluf, A., and Liu, J. (1996). User-expertise modeling with empirically derived probabilistic implication networks. *User Modeling and User-Adapted Interaction* 5, 283–315.

[7] Desmarais, M.C., Xu, P. and Beheshti, B. (2015). A partition tree approach to combine techniques to refine item to skills Q-matrices *8th Conference on Educational Data Data Mining* (EDM 2015), Madrid, Spain.

[8] Osborne, Jason W. & Amy Overbay (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation*, 9(6). Retrieved October 31, 2015 from http://PAREonline.net/getvn.asp?v=9&n=6

[9] Pavlik Jr., P.I., Cen, H., Wu, L., Koedinger, K.R.: Using Item-type Performance Covariance to Improve the Skill Model of an Existing Tutor. In: Baker, R.S., Beck, J.E. (Eds.) *Proceedings of the 1st International Conference on Educational Data Mining*, 2008. Montreal, Canada, p. 77-86

[10] Pavlik, P.I., Cen, H., Koedinger, K.R.: Learning factors transfer analysis: using learning curve analysis to automatically generate domain models. In: Barnes, T., Desmarais, M.C., Romero, C., Ventura, S. (eds.) 2nd

International Conference on Educational Data Mining--EDM2009, pp. 121-130, Cordoba, Spain, 1-3 July 2009

[11] Scheines, R., Silver, E., Goldin, I.: Discovering prerequisite relationships among knowledge components. In: Stamper, J., Pardos, Z., Mavrikis, M., McLaren, B. (eds.) *Proceedings of the 7th* International *Conference on Educational Data Mining*. pp. 355–356. European Language Resources Association (ELRA), May 2014

[12] Tatsuoka, K K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement. 20*(4) 345-354.