

A Method for Finding Prerequisites Within a Curriculum

Annalies Vuong, Tristan Nixon, and Brendon Towle¹, Carnegie Learning

Creating an educational curriculum is a difficult task involving many variables and constraints [Wang 2005]. In any curriculum, the order of the instructional units is partly based on which units teach prerequisite knowledge for later units. Historically, psychologists and cognitive scientists have studied the dependency structure of information in various domains [Bergan and Jeska 1980; Griffiths and Grant 1985; Chi and Koeske 1983]; however, many of these studies have been hampered by statistical issues such as the difficulty of removing instructional effects when using small samples [Horne 1983]. We hypothesize that large-scale assessment data can be analyzed to determine the dependency relationships between units in a curriculum. This structure could then be used in generating and evaluating alternate unit sequences to test whether omitting or re-ordering units would undermine necessary foundational knowledge building. Our method incorporates all possible pair-wise dependency relationships in a curriculum and, for each such candidate dependency, compares performance of students who used the potential prerequisite unit to performance of students who did not. We implemented the method on a random sample of schools from across the U.S. that use Carnegie Learning's Cognitive Tutor software and its associated curricula; a sample that far exceeds those used in previous studies both in size and in scope. The resulting structure is compared to a pre-existing list of prerequisites created by Carnegie Learning based on student skill models. We discuss extensions of this method, issues in interpreting the results, and possible applications. We hope that this work serves as a step toward developing a data-driven model of curriculum design.

Additional Key Words and Phrases: data mining, curriculum, intelligent tutor

1. INTRODUCTION

Curriculum is a fundamental part of education at every scale, from a one-day class to a four-year degree. Designing a curriculum involves balancing many competing constraints, not least of which is prerequisite knowledge. The method we present applies to any scale of curricula, though our data comes from mathematics curricula spanning one school year.

Prerequisite knowledge is here defined as the skills and information necessary to succeed in a given instructional unit within a curriculum. This knowledge can be acquired inside or outside the curriculum, giving rise to three important questions. What prerequisite knowledge is required to successfully learn each topic in the curriculum? Which units in the curriculum teach this knowledge? Finally, have students already acquired this knowledge outside of the curriculum? The third question will have an answer unique to each instructional situation. However, the first and second questions depend only on the content of the curriculum, and we believe that they can be answered empirically.

In fact, we believe they can be combined into one question. Given an instructional unit – call it Unit B – which prior units significantly influence students' success in this unit? It seems reasonable to conclude that these prior units cover some prerequisite knowledge. This paper lays out a method, given sufficient user data, for finding the prerequisite units for each instructional unit in a curriculum.

Creating a curriculum always involves defining prerequisites implicitly or explicitly. However, these definitions are usually based on expert opinion or on a theoretical model [Bergan and Jeska 1980; Griffiths and Grant 1985; Chi and Koeske 1983] and are rarely tested empirically. Even algorithms for designing optimal curricula may take prerequisite relations as a given [Wang 2005]. This lack of empirical testing is understandable, as it can be difficult to assess causal relationships or remove instruc-

¹Authors' addresses: A. Vuong, T. Nixon, and B. Towle, Carnegie Learning, Inc., 437 Grant St., Pittsburgh, PA 15219. Correspondence can be sent to avuong@carnegielearning.com.

tional effects, especially in small observational studies [Horne 1983]. Using a small sample also runs the risk of only answering the third question: do these students already have sufficient prior knowledge. A further difficulty in determining dependency structure is that a curriculum of multiple units will have a significant number of possible prerequisites to test.

Many of these problems relate to having small sample sizes. Our data are collected from students using Carnegie Learning's Cognitive Tutor, by far the most widely used intelligent tutoring system in the United States: it is currently used by over five hundred thousand students in more than twenty-five hundred schools. For high school mathematics, Carnegie Learning offers four Cognitive Tutor curricula. Each of the curricula consists of a sequence of units; each unit consists of a sequence of sections. Units cover distinct mathematical topics; sections cover distinct sets of problems on that topic, with a distinct student skill model for each section. Teachers and school administrators can create customized variations of the standard curricula by omitting units, reordering units, or adding units from another curriculum, though they cannot customize the skill models or problem sets within each unit. In school year 2008-2009, 85% of teachers used a customized curriculum. The great variety of curricula variations thus created provided us with a natural opportunity to compare student performance as they progressed through subtly different unit sequences.

That all students used the same software was very helpful in avoiding issues raised in previous studies: it reduced instructional effects in the data, reduced the chance of content-based false positives via the coverage of a large number of topics, and provided sufficiently uniform data to test almost all possible prerequisites.

2. METHOD

2.1. Data

Our data is taken from a random one-fifth sample of all schools using Cognitive Tutor in the 2008-2009 school year from whom were collected detailed logs of students' activity in the software. This sample comprises 20,577 students from 888 schools across the United States.

The standard Carnegie Learning high school curricula – Bridge to Algebra, Algebra I, Geometry, and Algebra II – contain 175 total units, including cross-listed units. Every unit prior to a given unit within each curriculum was considered a possible prerequisite; “true” prerequisites of a unit are defined to be those which have a significant effect on the success of students in completing the target unit. The list of all possible prerequisite relationships can be represented as a list of pairs of the form (Unit A, Unit B), where A is prior to B in one of the four curricula. For each pair of units, Sample A comprised all students in the data set whose curricula included both Unit A and Unit B and who attempted at least one problem in both units. Sample B contained all students whose curricula included Unit B but omitted Unit A and who attempted at least one problem in Unit B. We tested all such pairs for which there was at least one student in each sample, resulting in 3,832 tested pairs. Lacking enough information to compare a pair of units was rare: only 30 pairs were not tested. The average Sample A size was 325.5 students, with average Sample B size at 506.8 students.

2.2. Testing

For any given Unit B, student performance in the first section of the unit is more likely to be affected by prior knowledge from other units than performance in the later sections, since later sections would likely rely strongly on knowledge from earlier in the unit. To avoid this confound, we decided to evaluate success in Unit B by looking at student performance on the first section of the unit. Our hypothesis was that if

Unit A truly provides prerequisite knowledge for Unit B, we should see an increased graduation rate on the first section of Unit B for students who have had some exposure to the material in Unit A. We therefore calculated the overall graduation rate from the first section of Unit B in each of the two samples.

In the Cognitive Tutor, a student graduates from a section only if they master all of the section's skills over the course of a reasonable number of problems. The system will automatically promote them to the next section if they hit the problem limit without mastering all skills. Additionally, a student may simply stop working on a section and never return, for reasons such as leaving the class. Finally, a teacher may move a student forward out of a section if they are not keeping pace with the rest of class. Given the ways in which a student may leave a section without graduating, it is reasonable to assume that in general the performance of students who graduated from the first section of Unit B was better than the performance of those who did not and thus reasonable to use graduation rate as a performance metric.

To compare graduation rates for each pair of units in our list of possible prerequisites, we used a binomial test with $\alpha = 0.01$. The binomial test for each pair looked for a significant difference in the average graduation rate for Sample A as compared to the rate for Sample B. If a significant difference was found, Unit A was deemed a true prerequisite for Unit B.

3. RESULTS

The average number of true prerequisites for each unit was approximately 9.6 out of an average of 21.2 possible prerequisites. Overall, a little over 43% of all possible prerequisites were found to be true prerequisites.

As part of our analysis, we compared the data-driven prerequisites to the list of prerequisite relationships which is already included in the Cognitive Tutor, a list generated primarily from shared skills in the cognitive models for different units. Figure 1 shows all possible (Unit A, Unit B) pairs across the four curricula as elements in a 175 by 175 matrix, with Unit A as the row and Unit B as the column. The order of the units within each curriculum and the order of curricula (Bridge to Algebra, Algebra I, Geometry, Algebra II) was preserved in each axis of the table, hence the triangular regions indicate the standard curricula. In Figure 1, red indicates a non-significant relationship between Units A and B, with green indicating a significant relationship. Colored blocks outside the triangular regions display the results of cross-listed units, and white blocks indicate inter-curricular pairs not on our list of possible prerequisites.² Yellow blocks mark the small number of cases where there was insufficient data.

Figure 2 shows the difference between the set of empirically-derived prerequisite relationships and the set of prerequisite relationships given by the Cognitive Tutor for the same list of possible prerequisites. Dark green indicates a true prerequisite relationship found in both sets; light green a relationship deemed not prerequisite by both. Orange indicates a prerequisite relationship listed by the Cognitive Tutor that was not found in the data. Yellow shows a relationship found in the data but not in the Cognitive Tutor set. Overall, there was 56% agreement between the two methods (percentage of green blocks) and 14% agreement on true prerequisites (percentage of dark green blocks). The next section gives possible reasons for these differences.

4. DISCUSSION

Our goal was to devise a method which would empirically determine the prerequisite relationships in a given curriculum. It is important to note that this method is dis-

²Units were compared only within each curriculum. A study allowing for cross-curricular prerequisites could yield further pertinent results.

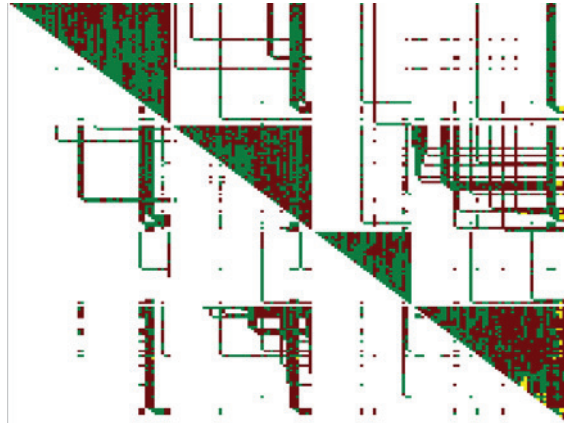


Fig. 1. Heatmap of data-based prerequisite relationships. Red indicates a non-significant relationship between Units A and B; green, a significant relationship; yellow, insufficient data to compare; and white, a pair not on our list of possible prerequisites.

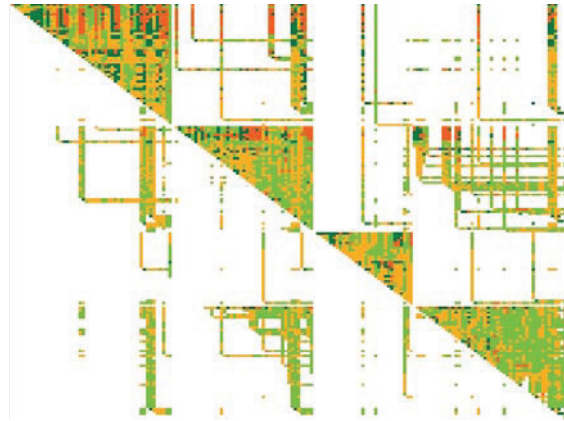


Fig. 2. Heatmap comparing data-based and cognitive model-based prerequisite relationships. Dark green indicates a true prerequisite relationship found in both sets; light green, a relationship deemed not prerequisite by both; orange, a prerequisite relationship listed by the Cognitive Tutor that was not found in the data; and yellow, a relationship found in the data but not in the Cognitive Tutor.

tinctly different from methods to derive student skill models from data (e.g. [Cen et al. 2006; Barnes et al. 2005]). We feel that our method is complimentary to such work. Our analysis takes place at a larger grain-size, comparing the relationships between units of instruction that consist of distinct skill models. The similarity or difference between these skill models may play an important role in the prerequisite relationship between the units.

It is interesting to note how the empirically derived prerequisite structure deviates from that determined by domain experts. Where the empirical evidence shows a prerequisite not identified by a domain expert, we can imagine mathematical skills practiced in both units but not part of the focus for the units, and thus easily overlooked in the expert analysis. Expert judgment is more suspect when it has identified a prerequisite that is not borne out by the data (marked in orange). However, as we see from Figure 2, this type of disagreement between expert and empirical data was confined mainly to the earlier units in each standard curriculum. It may be that a large

number of students were already sufficiently proficient on this preliminary material, to the point where any practice on prerequisite material would make no appreciable improvement to their performance. In this case, these could be true prerequisites but for our population it would make more sense to treat them as false prerequisites.

These results do raise many new questions, especially where this analysis disagrees with expert assessment, and further investigation of such units is warranted. There are many possible hypotheses for any given pre-requisite relationship. It could be that the skills practiced in unit A are truly foundational for those in B. Conversely, it could also be the case that the units share some subset of skills in common, and that it is simply the previous practice with these skills which provides the improved performance on unit B. In such a case, the curriculum structure is effectively loading the learning of those common skills onto whichever unit occurs first. It is also possible that another metric of student performance might reveal a difference not revealed by graduation rates. Although our method does not answer such questions, it does provide a useful framework for focusing attention to those unit pairs which deserve more investigation.

Since there is not already an agreed-upon way to objectively determine prerequisites, it is difficult to assess the validity of our method. A possible assessment method would be to see if different measurements of student performance yield the same prerequisite relationships. One such measure of performance is the number of problems done by students before graduating. Preliminary data exploration suggests that the result with this metric will be similar.

The dependency structure of a curriculum can be a powerful piece of information. For instance, [Ohland et al. 2004] found that removing a gateway course for their engineering major improved graduation rates for the major as a whole, suggesting that the course was not a true prerequisite. Every curriculum is subject to time constraints, and knowing which units can be safely omitted if necessary is important. Given this information, we could help teachers to customize their curricula without removing necessary prerequisite units.

Overall we feel that data-driven course design is a fruitful topic for research, which could yield relevant and useful information in many different areas of education.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Steve Ritter of Carnegie Learning for his guidance and advice.

REFERENCES

- BARNES, T., BITZER, D., AND VOUK, M. 2005. Experimental analysis of the Q-matrix method in knowledge discovery. *Foundations of Intelligent Systems*, 603–611.
- BERGAN, J. AND JESKA, P. 1980. An examination of prerequisite relations, positive transfer among learning tasks, and variations in instruction for a seriation hierarchy. *Contemporary Educational Psychology* 5, 3, 203–215.
- CEN, H., KOEDINGER, K., AND JUNKER, B. 2006. Learning Factors Analysis—A general method for cognitive model evaluation and improvement. *Intelligent Tutoring Systems*, 164–175.
- CHI, M. AND KOESKE, R. 1983. Network representation of a child's dinosaur knowledge. *Developmental Psychology* 19, 1, 29–39.
- GRIFFITHS, A. AND GRANT, B. 1985. High school students' understanding of food webs: Identification of a learning hierarchy and related misconceptions. *Journal of Research in Science Teaching* 22, 5, 421–436.
- HORNE, S. 1983. Learning Hierarchies: a critique. *Educational Psychology* 3, 1, 63–77.
- OHLAND, M., YUHASZ, A., AND SILL, B. 2004. Identifying and removing a calculus prerequisite as a bottleneck in Clemson's General Engineering curriculum. *JOURNAL OF ENGINEERING EDUCATION-WASHINGTON-* 93, 253–258.
- WANG, Y. 2005. A GA-based methodology to determine an optimal curriculum for schools. *Expert Systems with Applications* 28, 1, 163–174.