A Data-Driven Framework of Modeling Skill Combinations for Deeper Knowledge Tracing

Yun Huang
Intelligent Systems Program
University of Pittsburgh
Pittsburgh, PA, USA
yuh43@pitt.edu

Julio D. Guerra-Hollstein
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA, USA
jdg60@pitt.edu

Peter Brusilovsky
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA, USA
peterb@pitt.edu

ABSTRACT
This paper explores the problem of modeling student knowledge in complex learning activities where multiple skills are required at the same time and combinations of skills might carry extra specific knowledge. We argue that in such cases mastery should be asserted only when a student can fluently apply skills in combination with other skills. We propose a data-driven framework to model skill combinations for tracing students' deeper knowledge, and also propose a novel evaluation framework which primarily focuses on the mastery inference quality. Our experiments on two real-world datasets show that proposed model significantly increases mastery inference accuracy and more reasonably distributes students' efforts comparing with traditional Knowledge Tracing models and its non-hierarchical counterparts.

Keywords
complex skill, multiple skill, composition effect, robust learning, deep learning, Knowledge Tracing, Bayesian Network

1. INTRODUCTION
Knowledge Tracing (KT) [2] has been established as an efficient approach to model student skill acquisition in intelligent tutoring systems. The essence of this approach is to decompose overall domain knowledge into elementary skills and map each step's performance to the knowledge level of a single skill. However, KT assumes skill independence in problems that involve multiple skills, and it is not always clear how to decompose overall domain knowledge. Recent research demonstrated that there is additional knowledge related to specific skill combinations; in other words, the knowledge about a set of skills is greater than the "sum" of the knowledge of individual skills [6], some skill must be integrated (or connected) with other skills to produce behavior [9]. For example, students were found to be significantly worse at translating two-step algebra story problems into expressions (e.g., 800-40x) than they were at translating two closely matched one-step problems (with answers 800-y and 40x) [6]. In particular, research on computer science education has long argued that knowledge of a programming language cannot be reduced to a sum of knowledge about different constructs since there are many stable combinations (patterns, schemas, or plans) that have to be taught. We present a data-driven framework for modeling skill combinations and evaluating student models for adaptive tutoring in order to achieve deeper knowledge tracing.

2. PROPOSED FRAMEWORK

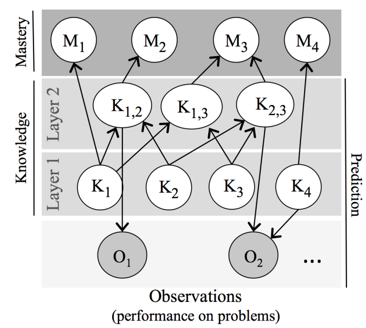

Figure 1: The Bayesian network structure of CKM-HSC.

We construct a Bayesian network called *conjunctive knowledge modeling with hierarchical skill combinations (CKM-HSC)* with the following knowledge structure:

I The first layer consists of basic individual skills (e.g., K_1) that capture the basic understanding of each skill.

II The intermediate layers consist of skill combinations (e.g., $K_{1,2}$), which can be derived from smaller skill units that capture a deeper knowledge level of each individual skill. Now, we consider only skill combinations from two basic individual skills.

III The last layer consists of *Mastery* nodes (e.g., M_1) for each individual skill, which reflects the idea of granting a skill's mastery based on relevant skill combinations' knowledge levels. Now, we compute the joint probability of each relevant skill combination being known as the probability of the current skill being mastered.

To learn the network structure, we propose a greedy search algorithm where a pre-ordering of the skill combination candidates is given as input, and during each iteration, the data likelihood of the network incorporating a new skill combination is compared to that of the optimal network so far. We now replace the search procedure with an empirical thresholding method, which generates an almost identical network with much less time. It selects combinations based on the following criteria: 1) the difficulty difference between the combined skill and its hardest individual one should be positive and large; 2) the difficulty of the combined skills should be high; 3) an item with higher difficulty should be more likely to require combined skills; and 4) each item can only have a limited number of skill combinations. To perform a dynamic knowledge estimation, we use the roll-up mechanism, as in [1]. For performance prediction, we apply Noisy-and gates on item nodes (e.g., O_1) as in [1, 3].

Table 1: Dataset descriptive statistics.

Dataset	#obs.	#items	#skills	avg #skills/item	#users	%correct
SQL	17,197	45	34	5 (from 1 to 10)	366	58%
Java	25,988	45	56	5 (from 1 to 11)	347	67%

To address the limitation of predictive performance metrics [7, 5], we propose a multifaceted data-driven evaluation framework that includes mastery accuracy and effort, the item discriminative index [3], and performance prediction metrics. The basic idea of the mastery accuracy metric is that once a student model asserts mastery for an item's required skills, the student should be unlikely to fail the current item. Meanwhile, the mastery effort metric empirically quantifies the number of practices that are needed to reach a level of mastery for a given set of skills. These metrics extend our recent learner effort-outcome paradigm [5] and Polygon multifaceted evaluation framework [7].

3. STUDIES

We used datasets collected from SQL and Java programming learning systems from 2013 to 2015 at the University of Pittsburgh. Table 1 shows the descriptive statistics (with multiple attempts). We conducted a 10-fold student stratified cross-validation. For each metric, we reported the average value across 10 folds and with a 95% confidence interval, based on the t-distribution. We used the Bayes Net Toolbox to construct all the models. On average, we extracted 14 and 30 skill combinations on SQL and Java datasets.

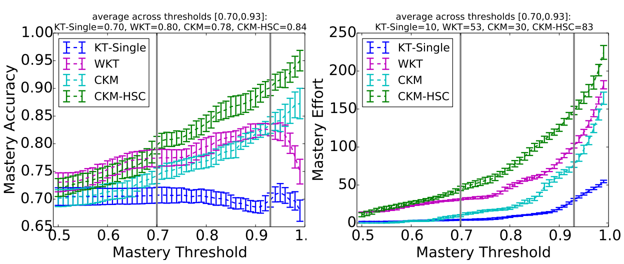

Figure 2: Mastery accuracy and effort comparison on Java dataset. Grey lines denote regions with enough data points to compute mastery metrics and with high enough values to be considered as proper mastery thresholds.

Our first study investigates whether the proposed skill combination incorporated model is better than traditional KT models. We compare classic Knowledge Tracing (KT-Single) [2], Weakest Knowledge Tracing (WKT) [4], and our proposed conjunctive knowledge modeling without (CKM) or with skill combinations (CKM-HSC) (Figure 2). On both datasets, CKM-HSC has a comparable predictive performance to other models, but it has significantly better mastery accuracy than other models. Although it requires more efforts to reach mastery, we think that such "extra" practices is necessary for reaching an acceptable mastery inference accuracy. We further conduct a drill-down analysis for mastery effort by splitting skills into two groups based on whether they involve skill combinations. We find out that for skills that involve skill combinations, WKT would blindly distribute students' efforts among different application contexts, risk students reaching mastery by practicing simple problems, and also guide students to spending more efforts on skills without combinations. On the other hand, CKM-HSC saves students' efforts on basic individual skill understanding and on skills without skill combinations. It requires students to focus more on applying skills in different contexts combined with other skills. We further conduct two studies demonstrating that using a hierarchical structure is better than using a flat independent structure for incorporating skill combinations, and that our modeling can be improved by adding external knowledge (such as expert knowledge or skill combinations' textual proximity) for skill combination extraction. Details are reported in [8].

4. CONCLUSIONS

Our work serves as a first attempt to consider the skill application context for modeling deeper knowledge in a student model using data-driven techniques. We also propose a novel data-driven evaluation framework for such complex skill student models. We only consider pairwise skill combinations as the skill application context; it will be to interesting to consider more complex skill combinations. Such combinations should have a natural connection with the concept of *chunk* in cognitive psychology for defining expertise. Meanwhile, to address the problem of computational complexity we now employ some heuristics. We should explore alternative approaches and more efficient techniques. We will also consider working with larger datasets and datasets with more sparse connections among variables. We expect that our model can provide more benefits when deployed in real-world tutoring systems. For example, it might enable better remediation and raise students' awareness of pursuing true mastery.

5. REFERENCES

[1] C. Conati, A. Gertner, and K. Vanlehn. Using bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interaction*, 12(4):371–417, 2002.

[2] A. T. Corbett and J. R. Anderson. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 4(4):253–278, 1995.

[3] J. De La Torre. An empirically based method of q-matrix validation for the dina model: Development and applications. *Journal of educational measurement*, 45(4):343–362, 2008.

[4] Y. Gong, J. E. Beck, and N. T. Heffernan. Comparing knowledge tracing and performance factor analysis by using multiple model fitting procedures. In *Proc. 10th Int. Conf. Intelligent Tutoring Systems*, pages 35–44. Springer, 2010.

[5] J. P. González-Brenes and Y. Huang. Your model is predictive but is it useful? theoretical and empirical considerations of a new paradigm for adaptive tutoring evaluation. In *Proc. 8th Intl. Conf. Educational Data Mining*, pages 187–194, 2015.

[6] N. T. Heffernan and K. R. Koedinger. The composition effect in symbolizing: The role of symbol production vs. text comprehension. In *Proc. 19th Annual Conf. Cognitive Science Society*, pages 307–312.

[7] Y. Huang, J. P. González-Brenes, R. Kumar, and P. Brusilovsky. A framework for multifaceted evaluation of student models. In *Proc. 8th Int. Conf. Educational Data Mining*, pages 203–210, 2015.

[8] Y. Huang, J. Guerra, and P. Brusilovsky. Modeling skill combination patterns for deeper knowledge tracing. In *6th Int. Workshop on Personalization Approaches in Learning Environments (In Submission)*, 2016.

[9] K. R. Koedinger, A. T. Corbett, and C. Perfetti. The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, 36(5):757–798, 2012.