Automatic Adaptive Sequencing in a Webgame

Tong Mu[1], Shuhan Wang[2], Erik Andersen[2], and Emma Brunskill[1]

[1] Stanford University,{tongm, ebrun}@cs.stanford.edu
[2] Cornell University,{e.andersen.acad, sonatithew}@gmail.com

Abstract. Intelligent tutoring systems can improve student outcomes, but developing such systems typically requires significant expertise or prior data of students using the system. In this work we propose a new approach for automatically adaptively sequencing practice activities for an individual student. Our approach builds on progress for automatically constructing curriculum graphs and advancing a student through a graph using a multi-armed bandit algorithm. These approaches have relatively few hyperparameters and are designed to work well given limited or no prior data. We evaluate our method, which can be applied to a diverse range of domains, in our online game for basic Korean language learning and found promising initial results. Compared to an expert-designed fixed ordering, our adaptive algorithm had a statistically significant positive effect on a learning efficiency metric defined using in game performance.

Keywords: Adaptive Sequencing · Automatic Curriculum Generation · Educational Games

1 Introduction

As educational technology continues to grow in popularity, it is desirable to have scalable, robust methods for creating efficient and adaptive learning pathways for new tutoring systems. When no prior data is available, one potential way to ensure that a new adaptive tutoring system will yield strong learning benefits is to heavily involve experts to create a curriculum structure and specify model parameters of an assumed student model. However this type of approach can be time consuming and difficult to scale. To lighten this burden of human expertise, data driven approaches use collected data to create or refine statistical models of student learning (e.g. [6,9,15]). Such work has demonstrated significant improvements in student learning outcomes and/or learning efficiency. However this approach still requires either a significant amount of expert input to ensure the initial learning pathways (when only a few students have used the system) are beneficial, or may have significantly suboptimal performance for students early on.

In our work we propose a new method that, given a set of activities, will automatically create adaptive learning pathways for students without requiring prior data. Our approach builds on several developments: methods that take a set of skills labeled with features and automatically constructs a knowledge (prerequisites) graph among this set [2,21], and the work to use multi-armed bandits to progress through such

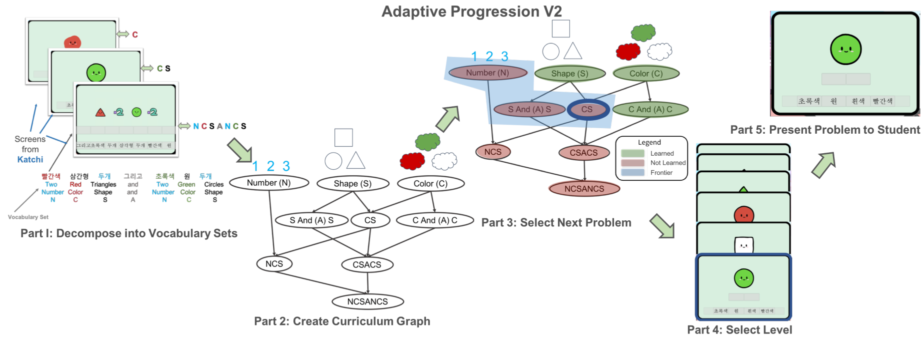

Fig. 1: Full pipeline of our adaptive algorithm applied to Katchi. Part 1 shows the grouping of individual items (such as Red and Green) into concepts (such as Color) and grouping problems into nodes, Part 2 shows organizing these nodes into a curriculum graph, Part 3 shows the adaptive algorithm's internal belief state of the mastery of each node, Part 4 shows selecting a specific level from within the selected node, and Part 5 shows presenting the selected problem to the student.

a graph [11]. This bandit based method only has a few hyperparameters, reducing the burden of expert time needed, and is demonstrated to potentially be more robust to variability in the student learning process [10].

Our method can be used in any domain where solving a pedagogical activity can be described by a simple program or by the execution trace of that program, which has been shown to cover diverse domains including basic arithmetic [2], logical proofs [1], and algebra [16]. We experimentally investigate our method in a Korean language learning webgame and perform analysis using evaluation metrics based on in game performance. We found that while there was no significant difference in total learning, learners in the adaptive progression condition had a statistically significant higher learning efficiency compared to students in the expert-designed fixed progression condition. This initial finding highlights the potential of our method to increase learning efficiency and save students time.

Related Work Our work considers the question of adaptively sequencing educational content. Bayesian Knowledge Tracing (BKT) [12] can be used to adaptively monitor a student's learning given a decomposition of problems into knowledge components (KCs), but does not specify how to select amongst the unknown skills.

Various recent approaches [13,19,5,15] often found promising results. However our work differs in that these previous works all use prior student data to create an adaptive algorithm. Most recently, Bassen et al. [4] used deep reinforcement learning (RL) with minimal expert input and no prior data to learn to sequence problems efficiently. They demonstrate good performance for learners after the experience from the first 200 learners. In general, even the most efficient deep RL methods will require at least such an amount of experience to achieve good performance. Their method does not discuss a way to enforce a curriculum graph for the initial learners when concepts and activities exhibit strong prerequisite dependencies as they do in our domain. Methods such as ours that can enforce a curriculum can improve the experience for early learners.

Our work builds on the adaptive algorithm presented by Clement et al. that automatically provides personalized student advancement through curriculum graphs using concepts of a ZPD and multi-armed bandits [11]. This work previously assumes an expert specified curriculum graph was given, which we do not. In our work, we automatically generate a curriculum graph from a set of practice activities using developments in automatic curriculum generation using execution traces [2,20]. The use of the curricula generated by these automatic methods for adaptive sequencing have not been previously investigated. As a result, our work aims to provide, to our knowledge, one of the first evaluations of a system that, using no prior data and very little expert input, automatically creates a curriculum graph and an adaptive progression for students from a set of practice items.

2 Method

Our method is illustrated in Figure 1 and described in more detail below.

2.1 Domain: Katchi, a Korean Language Learning Webgame

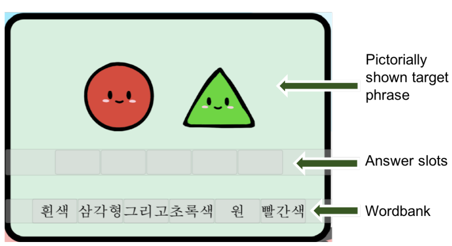

Fig. 2: A Screenshot of the *Red Circle And Green Triangle* activity of Katchi. learners drag and drop words from the wordbank at the bottom of the screen to the answer slots to describe the pictorially shown target phrase.

We run our experiments on our online Korean learning webgame *Katchi*. The goal of Katchi is for students to master simple Korean phrases of the form *Number Color Shape And Number Color Shape* for all possible combinations of the taught vocabulary. To pass an activity (Fig 2), learners must drag and drop the Korean words from the wordbank into the correct answer slots to describe the pictorially shown target phrase. Activities vary in complexity in terms of the number of words needed, ranging from having 1 to 7 empty slots. Learners are allowed an unlimited amount of attempts on each activity and must answer the activity correctly to advance. An activity bank containing 402 unique activities was created.

2.2 Automated Curriculum Graph Creation:

We first build a hierarchical structure to model difficulty dependencies between a set of pedagogical items. We follow Wang et. al's [2,21] work in automatic curriculum generation for language learning and define:

Definition 1 ([21]). *A phrase* s_1 *is* harder than *another phrase* s_2*, indicated as* $s_1 > s_2$*, if* s_1 *is longer than and covers all the vocabulary words in* s_2*. A phrase* s_1 *is* directly harder than s_2 *if there does not exist a third phrase such that* $s_1 > s_3 > s_2$*.*

This definition implies a directed graph where a directed edge represents a "directly harder than" relation between two phrases. We represent each level that involves multiple words as a high level conceptual phrase (for example, a node is *color shape*, and the activities *red triangle* and *green square* will both fall under this node). We represent all single word levels as separate nodes to ensure students will learn each of the basic vocabulary words. The curriculum graph used can be seen in Part 2 of Figure 1.

While we evaluate our method in language learning, there exist similar curriculum generation methods that use solution execution traces, or the steps taken to solve an educational activity [2]. These methods can be applied to any domain where automatic solution generation is possible, which include a diverse range of domains such as arithmetic and logical proofs.

2.3 Automatic Progression through Curriculum Graph

Given a curriculum graph $\mathcal{G}$ over a set of nodes representing concepts and with one or more practice materials mapped to each node, our algorithm for selecting the next item to practice consists of combining a model of forgetting [17] and a model of the zone of proximal development (ZPD) [11]. We incorporate a model of forgetting as we consider a language learning setting where forgetting is known to be very important [3,7,17] and spaced repetition has long been a gold standard. Overall our method induces a policy that interleaves review activities with learning activities on which the student is making the fastest learning progress.

Progression Metric: To progress students to items farther in the curriculum, we need a signal of learning. Measuring the true knowledge state of students is a key challenge in the online game setting. As a proxy, we use correctness on the first attempt on an activity from the node to mark the node as learned.

Node selection: To select a node from a curriculum graph to present to the student, our method tracks three subsets of nodes for that student as shown in Figure 1 Part 3: the learned set ($\mathcal{L}$), the not learned set, and the frontier set ($\mathcal{ZPD}$). The frontier set consists of nodes in the not learned set on the boundary of the learned set. Such items are considered to be the Zone of Proximal Development (ZPD). Following prior work in psychology that hypothesized students will learn best when they are given activities that are at the appropriate level of difficulty[8], items in the ZPD are prime targets for learning. Upon initialization, all nodes with no incoming edges define the frontier and all nodes are in the not learned set. To select a node, the algorithm first checks if there are any review nodes from $\mathcal{L}$ that should be presented (described further below). If not, a learning node from $\mathcal{ZPD}$ is chosen following the bandit based ZPDES algorithm [11](See original paper for details). A node is moved to $\mathcal{L}$ once student performance on the node satisfies the mastery progression metric and an unlearned node is added to $\mathcal{ZPD}$ if all prerequisites of that node are marked as learned.

Review Nodes: To incorporate review, we track potential forgetting using the MCM model of forgetting [17]. MCM models the memory strength of items using a sum

of exponential decaying memory traces left each time an item is reviewed. A node is marked as needing review if its memory strength is lower than a threshold.
Selecting An Item From A Node: Once a node is selected, an activity from the node needs to be selected. Activities need to be selected in a way that ensures all basic components are practiced and students understand the concept generally taught by each node as opposed to only subsets of its instantiation. To give an extreme example of an undesired situation, if a student only sees *Circle* in problems that involve *Shapes* and never experiences *Square*, then a student who started out as a novice would not be able to complete problems involving *Square* even if they could complete problems of that node with *Circle*. It is infeasible to include all 402 unique activities so we instead ensure students see a varying array of vocabulary through time which the expert baseline was carefully designed to do. To ensure this in our adaptive algorithm, we use a second MCM model over each basic vocabulary word and at every timestep we choose the item from the node with the vocabulary word that has been practiced least recently.

3 Preliminary Experiment

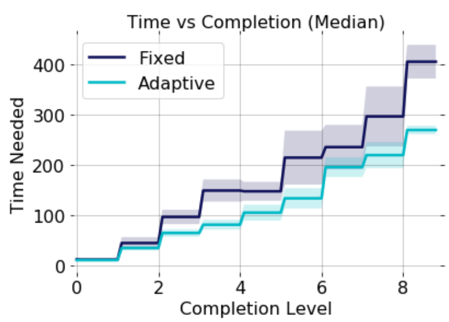

Fig. 3: Results from our experiment. Figure (a) shows completion level with time, with 95% confidence intervals shaded. We see the adaptive progression allows learners to reach completion levels with consistently less time than the fixed condition.

We compared our method, which we will refer to as the **Adaptive** progression, with an expert designed **Fixed** progression. The Fixed progression consists of 43 levels, 34 of which are unique, and was carefully created to achieve the learning goals of Katchi. We posted our game on a popular gaming website, Newgrounds[3] and uniformly at random assigned each learner to a condition. We collected data from 117 and 89 learners assigned to the Adaptive and Fixed conditions respectively.
Evaluation Metric: In this initial experiment, we focus on using within task signals for evaluation. Correctness on the first attempt is a strong signal of learning as for all activities that have more than 1 slot (which is 6 out of the 9 nodes, and 393 of the 402 possible unique activities) the probability of guess is very low. For the simplest multi-slot activity, the *Color Shape* activity, the probability of answering correctly through random guess is $\frac{1}{12}$. For the most complex, such as the activity *Two Green Circles And Three Red Triangles*, the probability of guess is $\frac{1}{5040}$. We

[3] `https://www.newgrounds.com/`

use this to define a completion metric which counts a node as **completed** once a learner answers a problem from that node correctly on the first attempt. We define the **completion level** of a learner at a given time as the number of unique nodes they have completed so far. We evaluate the overall learning efficiency in terms of number of completed nodes per minute (**CNPM**), of all the learners before dropout.
Results: We did not find a statistically significant difference in the average amount of completion upon dropout among conditions, meaning learners dropped out at similar stages of material difficulty (the mean completion level at dropout for Fixed and Adaptive were 2.6 and 2.4 respectively, a Mann-Whitney-U test results in $p = 0.37$ ($U = 2457$)). However we did find a statistically significant difference in learning efficiency. We found the adaptive algorithm overall enabled learners to progress through the same material faster than the fixed progression. On average, learners in the adaptive progression progressed at 1.7 CNPM while learners in the fixed progression progressed at 1.2 CNPM. Figure 3 shows completion through time. Due to differential dropout, there is overlap at different points, but overall the adaptive progression is progressing learners faster. Accounting for multiple comparisons using the Bonferroni Correction, the result of a Mann-Whitney-U test suggested there was a statistically significant difference between the conditions at the $\alpha = 0.01$ level, ($U = 1677$, $p = 0.003$). Increases of learning efficiency are very benefitial as it allows learners to save time and mental energy to put towards other studies. This is especially so as there has been research that shows faster and slower learners that learn a concept to the same level of mastery, irregardless of the amount of practice items needed to reach that level, showed the same strength of learning in terms of rate of forgetting [18,14,6]. Therefore our initial findings which suggests our method can potentially increase learning efficiency is promising.

4 Conclusion

Following evidence that creating personalized and adaptive educational systems can lead to improved learning, our work presented a novel system that can take in practice items and descriptions of those items in terms of its underlying skills, and automatically create an adaptive personalized sequence of the material for a student. Key features of our adaptive algorithm is that it only needs minimal expert input and does not require existing student data to set the hyperparameters of adaptivity. Our method can be applied to a wide range of domains and we run a preliminary study in a language learning domain to examine the effectiveness of our method. We found initial evidence our method may able to increase learning efficiency compared to a strong fixed progression.

5 Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. IIS-1657176 and the BIGDATA award. It is also supported by the Stanford Human Centered AI HoffmanYee grant and the Graduate Fellowships for STEM Diversity. Additionally, we thank Brandon Cohen, Nicholas Teo, Evan Adler, Urael Xu, and Nicole Cheung for their work in creating Katchi.

References

1. Ahmed, U.Z., Gulwani, S., Karkare, A.: Automatically generating problems and solutions for natural deduction. In: Twenty-Third International Joint Conference on Artificial Intelligence (2013)
2. Andersen, E., Gulwani, S., Popovic, Z.: A trace-based framework for analyzing and synthesizing educational progressions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 773–782. ACM (2013)
3. Bahrick, H.P., Bahrick, L.E., Bahrick, A.S., Bahrick, P.E.: Maintenance of foreign language vocabulary and the spacing effect. Psychological Science **4**(5), 316–321 (1993)
4. Bassen, J., Balaji, B., Schaarschmidt, M., Thille, C., Painter, J., Zimmaro, D., Games, A., Fast, E., Mitchell, J.C.: Reinforcement learning for the adaptive scheduling of educational activities. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–12 (2020)
5. Brinkhuis, M.J., Savi, A.O., Hofman, A.D., Coomans, F., van der Maas, H.L., Maris, G.: Learning as it happens: A decade of analyzing and shaping a large-scale online learning system. Journal of Learning Analytics **5**(2), 29–46 (2018)
6. Cen, H., Koedinger, K.R., Junker, B.: Is over practice necessary?-improving learning efficiency with the cognitive tutor through educational data mining. Frontiers in artificial intelligence and applications **158**, 511 (2007)
7. Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., Rohrer, D.: Distributed practice in verbal recall tasks: A review and quantitative synthesis. Psychological bulletin **132**(3), 354 (2006)
8. Chaiklin, S.: The zone of proximal development in vygotsky's analysis of learning and instruction. Vygotsky's educational theory in cultural context **1**, 39–64 (2003)
9. Chi, M., VanLehn, K., Litman, D., Jordan, P.: Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. User Modeling and User-Adapted Interaction **21**(1-2), 137–180 (2011)
10. Clement, B., Oudeyer, P.Y., Lopes, M.: A comparison of automatic teaching strategies for heterogeneous student populations. In: EDM 16-9th International Conference on Educational Data Mining (2016)
11. Clement, B., Roy, D., Oudeyer, P.Y., Lopes, M.: Multi-armed bandits for intelligent tutoring systems. JEDM-Journal of Educational Data Mining **7**(n) (2015)
12. Corbett, A.T., Anderson, J.R.: Knowledge tracing: Modeling the acquisition of procedural knowledge. User modeling and user-adapted interaction **4**(4), 253–278 (1994)
13. David, Y.B., Segal, A., Gal, Y.K.: Sequencing educational content in classrooms using bayesian knowledge tracing. In: Proceedings of the sixth international conference on Learning Analytics & Knowledge. pp. 354–363. ACM (2016)
14. Gentile, J.R., Voelkl, K.E., Pleasant, J.M., Monaco, N.M.: Recall after relearning by fast and slow learners. The Journal of experimental education **63**(3), 185–197 (1995)
15. Mandel, T., Liu, Y.E., Levine, S., Brunskill, E., Popovic, Z.: Offline policy evaluation across representations with applications to educational games. In: Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems. pp. 1077–1084. International Foundation for Autonomous Agents and Multiagent Systems (2014)
16. O'Rourke, E., Butler, E., Tolentino, A.D., Popović, Z.: Automatic generation of problems and explanations for an intelligent algebra tutor. In: International Conference on Artificial Intelligence in Education. pp. 383–395. Springer (2019)
17. Pashler, H., Cepeda, N., Lindsey, R.V., Vul, E., Mozer, M.C.: Predicting the optimal spacing of study: A multiscale context model of memory. In: Advances in neural information processing systems. pp. 1321–1329 (2009)

18. Underwood, B.J.: Speed of learning and amount retained: A consideration of methodology. Psychological Bulletin **51**(3), 276 (1954)
19. Vainas, O., Bar-Ilan, O., Ben-David, Y., Gilad-Bachrach, R., Lukin, G., Ronen, M., Shillo, R., Sitton, D.: E-gotsky: Sequencing content using the zone of proximal development. arXiv preprint arXiv:1904.12268 (2019)
20. Wang, S., Andersen, E.: Grammatical templates: Improving text difficulty evaluation for language learners. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. pp. 1692–1702 (2016)
21. Wang, S., He, F., Andersen, E.: A unified framework for knowledge assessment and progression analysis and design. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. pp. 937–948. ACM (2017)