Combining Difficulty Ranking with Multi-Armed Bandits to Sequence Educational Content

Avi Segal[1], Yossi Ben David[2], Joseph Jay Williams[3], Kobi Gal[1], Yaar Shalom[1]

[1] Ben-Gurion University of the Negev, Israel
[2] Microsoft, Israel
[3] National University of Singapore, Singapore

Abstract. We address the problem of how to personalize educational content to students in order to maximize their learning gains over time. We present a new computational approach to this problem called MAPLE (Multi-Armed Bandits based Personalization for Learning Environments) that combines difficulty ranking with multi-armed bandits. Given a set of target questions MAPLE estimates the expected learning gains for each question and uses an exploration-exploitation strategy to choose the next question to pose to the student. It maintains a personalized ranking over the difficulties of question in the target set and updates it in real-time according to students' progress. We show in simulations that MAPLE was able to improve students' learning gains compared to approaches that sequence questions in increasing level of difficulty, or rely on content experts. When implemented in a live e-learning system in the wild, MAPLE showed promising initial results.

1 Introduction

As e-learning systems become more prevalent they are accessed by students of varied backgrounds, learning styles and needs. There is thus a growing need for them to accommodate individual difference between students and adapt to their changing pedagogical needs over time.

We provide a novel algorithm for sequencing content in e-learning systems that combines offline learning from students' past interactions with an online exploration-exploitation approach in order to maximize students' learning gains. Our algorithm, called MAPLE (Multi-Armed Bandits based Personalization for Learning Environments), extends prior multi-armed bandit approaches in education [2], by explicitly considering question difficulty when initializing the online behavior of the algorithm, and when updating its behavior over time.

We first evaluated MAPLE in a simulation environment comparing its performance to a variety of sequencing algorithms. MAPLE outperformed all other approaches for average and strong students while showing the need for further tuning for weak students. We then implemented MAPLE in the wild in an existing e-learning system in a school with 7th grade students. MAPLE showed promising results when compared to an existing educational expert approach, and a state of the art approach based on Bayesian Knowledge Tracing [3]. Further experiments with larger student groups are needed.

2 Related Work

Our work relates to past research on using historical data to sequence content to students, and to work on multi-armed bandits for online adaptation of educational content.

Several approaches within the educational artificial intelligence community have used computational methods for sequencing content to students. Ben David et al. [3] developed a BKT based sequencing algorithm. Their algorithm (which we refer to in this paper as YBKT) uses knowledge tracing to model students' skill acquisition over time and sequence questions to students based on their mastery level and predicted performance. It was shown to enhance student learning beyond sequencing designed by pedagogical experts. Segal et al. [6] developed EduRank, a sequencing algorithm that combines collaborative filtering with social choice theory to produce personalized learning sequences for students. The algorithm constructs a difficulty ranking over questions for a target student by aggregating the ranking of similar students when sequencing educational content.

Multi-armed bandits provide a fundamental model for tackling the "exploration-exploitation" trade-off [7, 1]. Williams et al. [8] used Thompson Sampling to identify highly rated explanations for how to solve Math problems, and chose uniform priors on the quality of these explanations. Clement et al. [2] used human experts' knowledge to initialize a multi-armed bandit algorithm called EXP4, that discovered which activities were at the right level to push students' learning forward. In our work we do not rely on human experts, but rather use personalized difficulty rankings to guide the initial exploitation and update steps of our algorithm.

3 Problem Formulation and Approach

We consider an e-learning setting with a group of students S and a set of practice questions Q. The sequencing problem requires choosing at each time step a question to present to the student that will maximize her learning gains over the length of the practice session. The goal is to present students with challenging problems, while ensuring a high likelihood that they will be able to solve these problems.

Our approach to solve the problem, called MAPLE, maintains a belief distribution over expected learning gains to the student for solving each of the questions in Q. This distribution is initialized with a personalized difficulty ranking over the questions in Q. MAPLE samples the next question to the student from this distribution and updates it at each step given the student's performance on the question and its inferred difficulty to the student. When a student successfully solves a question, the distribution is adjusted to make harder questions more likely to be presented, and explore a broader range of questions. When a student fails to solve a question, it is adjusted to make easier questions more likely to be presented, and explore a narrower set of questions.

4 Simulations with Synthesized Data

We compared four different sequencing algorithms in a set of simulations: (1) The *MAPLE* approach which used the EduRank [6] algorithm for difficulty ranking. (2)

The *Ascending* approach sequenced questions according to an absolute difficulty ranking that was determined by pedagogical experts. (3) The *EduRank* approach sequenced questions from the easiest estimated question to the hardest estimated question per student. (4) The *Naive Maple* approach sequenced questions using the multi-armed bandit algorithm with random initialization.

We model questions in the simulation using a $\langle$skill,difficulty$\rangle$ pair; students are modeled as a vector of skill values for each question type. The probability that a student successfully solves a question is based on Item Response Theory [4] and on the difference between her skill level and the level of the question. Estimates of student's skill levels are increased and decreased based on successes and failures in question solving (respectively), proportional to the question difficulty.

Each algorithm was run with 1000 simulated students, and sequenced 200 questions for each student. Each question belonged to one of 10 skills (uniformly distributed) and to one of 5 difficulty levels (uniformly distributed). Students' initial competency levels in each skill were uniformly distributed between 0 (no skill knowledge) and 1 (full knowledge of skill). All algorithms had access to "historical" data generated by the simulation engine in a pre-simulation step, to build their internal models.

In simulations (Figure 1), MAPLE outperformed all other algorithms for strong and average students. For weak students the Ascending and Naive Maple approaches failed altogether. Both MAPLE and EduRank presented initial good progress but then experienced a decline in average skill level. This implies that MAPLE's adaptation scheme needs to be improved for this segment of students.

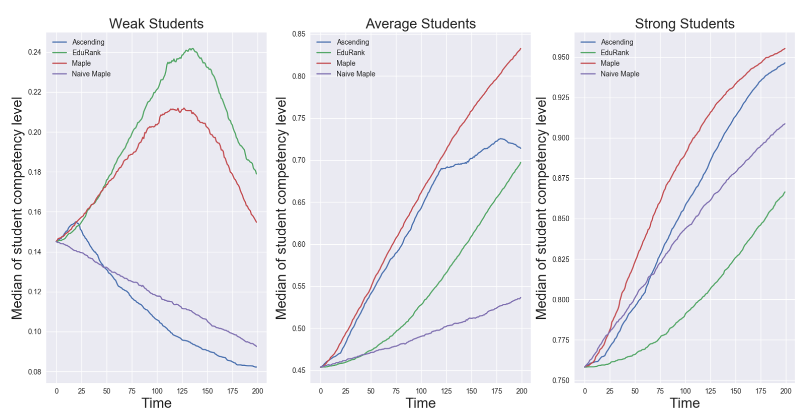

Fig. 1: Skill level progression per algorithm and student type.

5 Deployment and Evaluation in the Classroom

We then conducted a field study in the wild where students used different sequencing approaches in class. MAPLE was implemented in an e-learning system used for Math education. The study compared MAPLE in a school with 7th grade students to two existing sequencing algorithms. The experiment was conducted between May 9th 2017 and June 19th 2017 (end of school year). The students were randomly divided into 3 cohorts: (1) MAPLE Sequencing (2) YBKT Sequencing (3) Ascending Sequencing.

All students in the experiment were initially exposed to a pretest session. In this session they solved 10 questions hand picked by a pedagogical expert. Ninety two students solved the questions and there was no statistically significant difference between the three groups in the average score on this preliminary test. We thus concluded that each group exhibited similar baseline knowledge.

The students then engaged in multiple practice sessions in the e-learning system for the next 35 days, solving 10 assignment questions in each practice session. For each cohort, assignment questions were sequenced by the cohort's respective algorithm (i.e. MAPLE, YBKT or Ascending). At the end of this period, students were asked to complete a post test session, solving the same questions (in the same order) as in the pretest session. Twenty eight students completed the post test session. We attribute the decrease in students' response from pretest to post test to the pending end of the academic year (there was no difference in the dropout rates across the 3 cohorts).

Cohort	Time per Question (sec)	Average Grade
Ascending	6.49	43.76
MAPLE	10.69	71.28
YBKT	12.86	67.08

Table 1: Post test results per cohort: time per question and average grade.

Table 1 shows the students' average grade and the time spent on post-test questions. As can be seen, students assigned to the MAPLE condition achieved higher post test results than students assigned to the Ascending condition or to the YBKT condition. Further experiments with larger student groups are needed to evaluate statistical significance.

6 Conclusion

We presented a new method called MAPLE for sequencing questions to students in online educational systems. MAPLE combines difficulty ranking based on past student experiences with a multi-armed bandit approach using these difficulty rankings in online settings. We tested our approach in simulations and compared it to other algorithms. We then performed an initial experiment running MAPLE in a classroom in parallel to two baseline algorithms. MAPLE showed promising results in simulations and in the wild. Further simulations and experiments are needed to adapt and verify MAPLE's performance in more complex settings. For more information on the MAPLE algorithm and the simulation and field results please see [5].

References

1. Sébastien Bubeck, Nicolo Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.
2. Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, and Manuel Lopes. Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining (JEDM)*, 7(2):20–48, 2015.
3. Yossi Ben David, Avi Segal, and Ya'akov Kobi Gal. Sequencing educational content in classrooms using bayesian knowledge tracing. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, pages 354–363. ACM, 2016.
4. Ronald K Hambleton, Hariharan Swaminathan, and H Jane Rogers. *Fundamentals of item response theory*, volume 2. Sage, 1991.
5. Avi Segal, Yossi Ben David, Joseph Jay Williams, Kobi Gal, and Yaar Shalom. Combining difficulty ranking with multi-armed bandits to sequence educational content. *arXiv*, 2018.
6. Avi Segal, Ziv Katzir, Kobi Gal, Guy Shani, and Bracha Shapira. Edurank: A collaborative filtering approach to personalization in e-learning. In *Educational Data Mining 2014*, 2014.
7. William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.
8. Joseph Jay Williams, Juho Kim, Anna Rafferty, Samuel Maldonado, Krzysztof Z Gajos, Walter S Lasecki, and Neil Heffernan. Axis: Generating explanations at scale with learnersourcing and machine learning. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale*, pages 379–388. ACM, 2016.