**Background Chapter draft**

**Liu Rongxing s1810054**

**Introduction:**

The rapid development in E-learning has provided students with more flexible learning tools. Besides notes and lectures provided by schools, students can also choose a much wider variety of learning material online. On the other hand, choosing the most suitable learning methods has been a new challenge for students and teachers. Some students may spend unproportionate amount of time online yet still not being able to find useful and helpful material for their studies; some may spend a lot of time on practices yet not getting desirable grades for exams. This paper will review some recently established methods on difficulty ranking and sequencing educational content and discuss possible aspects for improvement.

**Literature Review:**

In 2014, Avi Segal et al. introduced a new algorithm called EduRank that combines collaborative filtering algorithms with social choice theory to personalize educational content for students [1]. Rather than sorting questions based on expected performance, which is prone to error, the algorithm infers a difficulty ranking directly over the questions for a target student. The algorithm was tested on two large data sets and evaluated using two ranking scoring metrics— NDPM and AP. Its performance was compared with other known methods such as EigenRank Algorithm [2], User Based Collaborative Filtering (UBCF), a matrix factorization method using SVD, an expert ranking (denoted by CER) and Topic-Based Ranker (TBR) [3]. EduRank achieved the best result on both datasets and both metrics. The execution of the algorithm was also not time consuming. However, the research left the “cold-start problem” [9] in e-learning system in which there is a need to sequence material for new students with little or no history in the system unresolved.

In 2019, Avi Segal et al. improved upon his previous work on EduRank [4]. The algorithm was extended to handle the “cold-start problem” by incorporating a prior score for every question in the training set by averaging over the scores of all students that solved this question in the training set. The results showed that the EduRank+Prior method beat the traditional EduRank algorithm, with noticeable differences up to the eighth week of training. Furthermore, the algorithm was also compared with ASC sequencing approach based on pedagogical experts on their performances in classrooms. The results showed that students using EduRank approach solved more difficult questions and achieved higher grades, and having to solve more difficult questions did not discourage students’ motivation to use the e-learning system. In addition, the research showed there is little agreement among students in difficulty rankings for the same set of questions, emphasizing the need for personalization in sequencing questions. However, the experiment overlooked the possibility that students can seek for assistance on a difficult question which makes their scores higher, and time spent lower. Solving a question with assistance does not improve as much as solving a question individually.

In 2016, Yossi Ben David et al. introduced a Bayesian Knowledge Tracing model for sequencing questions to students [5]. It is based on using knowledge tracing to model students' skill growth over time and to choose questions that improve their learning within the range of student's abilities, as determined by the model. The model was shown to outperform other models including the basic BKT model, KT-IDEM [6], COUNT and KT-IDEM-COUNT that did not consider partial credit scores, penalty on multiple attempts or integrating item difficulty. They also incorporated the model into a sequencing algorithm and deployed in classrooms. Again, it outperformed the ASC sequencing approach as shown by students’ scores and engagement. However, the research did not include pre and post tests to measure the effects of our algorithms on students’ performance and learning gains more accurately.

In 2018, Avi Segal et al. introduced 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 [7]. MAPLE was compared with three other sequencing algorithms, namely the Ascending approach that sequenced questions according to an absolute difficulty ranking that was determined by pedagogical experts, the EduRank approach that provided a personalized difficulty ranking over questions for each student, and the Naive Maple approach that sequenced questions using the multi-armed bandit algorithm with random weights initialization (without the EduRank based difficulty ranking component). The results showed that MAPLE adapted well to students in various skill competency levels (the weak students with level under 0.33, the average students with level between 0.33 and 0.67, and the strong students with level above 0.67), giving more questions close to their abilities. In addition, MAPLE gave the largest increase in median of student competency level for both average students and strong students. However, the models did not do well on weak students. This could be due to an overestimates of weak students’ learning rate.

In 2021, Kobi Gal et al. introduced a Reinforcement Learning agent for optimizing sequencing of learning materials [11]. The training was conducted offline using the EdNet dataset. They used a greedy iterative feature augmentation pipeline, with the Expected Cumulative Reward (ECR) metric for assessing policies, to determine the optimal feature representation. In the end, 4 extra features are taken (expl\_received, ssl, prev\_correct, av\_fam). The research also explored the methods’ robustness to perturbations in the environment induced by strong and weak learners and found that larger representations are more robust towards deviations from the expected dynamics derived from the data.

**Here is a brief summary of different work on sequencing and difficulty ranking:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tool | Year | Paper | Functionality | Problems/Notes |
| Difficulty Ranking Approach | 2019 | A Difficulty Ranking Approach to Personalization in E-Learning  <https://arxiv.org/abs/1907.12047> | EduRank constructs a difficulty ranking for each student by aggregating the rankings of similar students using different aspects of their performance on common questions. | The experiment assumed that the amount of time students spent on questions and the number of attempts on a question could directly reflect the difficulty of the questions to the students. The experiment was also conducted in an open environment that allowed students to look for assistance from teachers, classmates or Internet. However, such approach ignores the possibility that students might spend much shorter time on a supposedly difficult question as they would look for assistance on that question. Furthermore, they can “solve” the problem correctly on the first attempt. In this case the results might be inaccurate. |
| Bayesian Knowledge Tracing | 2016 | <https://dl.acm.org/doi/10.1145/2883851.2883885> | The algorithm is based on a Bayesian Knowledge Tracing (BKT) model that incorporates partial credit scores, reasoning about multiple attempts to solve problems, and integrating item difficulty. | “We capped the increase or de- crease in value to the posterior over the skill level at 0.1, which was based on showing examples to the pedagogical experts.” Since the most important idea of the project is about personalization, why can’t this value also be personalized?  As pointed out in the first paper, “indeed, it may occur in multiple choice questions that a student guessed the correct answer on the first attempt, even though the question was quite difficult.” This is quite common among students, especially nowadays when education tends to be more exam oriented. |
| Sequence Educational Content | 2018 | <https://arxiv.org/abs/1804.05212> | 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. | Unlike recommendation systems where students are free to choose from several questions, the idea of sequence educational content is more rigid. One or two questions given making the student bored (too easy) or frustrated (too difficult) will immediately discourage students’ motivation to using the system.  There is also a possibility that a student solves a question correctly with a completely wrong concept. In this case we cannot assume that the student’s learning must be improved.  As pointed out by the results, “both MAPLE and EduRank presented initial good progress but then experienced a decline in the estimated skill level.” This could be due to the overestimate of weak students’ learning ability. |
| Reinforcement Learning | 2021 | <https://www.dropbox.com/s/ph9rnwkjao72k60/RL_LAK_22%20%2814%29.pdf?dl=0> | This paper studies the use of Reinforcement Learning (RL) policies for optimizing the sequencing of learning materials to maximize learning as measured by expected future student performance. | There is little contribution to ECR for MDP\_3 and MDP\_4, with only around 1%. Is this worth the amount of time taken for computation? As pointed out in the paper, “our search procedure involves exhaustively looping through every remaining feature in the feature pool Ω to form the subsets”, besides the two methods mentioned in the paper, one can also simply set a threshold value below which the algorithm will terminate. |

**Conclusion:**

After the review of previous works, I plan to work on incorporating students’ learning activities to improve the difficulty ranking approach. This is feasible because we can obtain related data from EdNet [8]. EdNet is a large-scale hierarchical dataset of diverse student activities collected by Santa [10], a multi-platform self-study solution equipped with an artificial intelligence tutoring system. It offers the data in four different datasets named KT1, KT2, KT3 and KT4.

KT1 consists of students' question-solving logs, which is the most basic and fundamental information that can be used by various deep-learning knowledge tracing models such as Deep Knowledge Tracing and Self-Attentive Knowledge Tracing. It could provide the data required in previously mentioned research. EdNet-KT2, the simplest action-based dataset of EdNet, consists of the actions related to question-solving activities. EdNet-KT3 incorporates such learning activities by adding the following actions to the EdNet-KT2 dataset. Such actions can be utilized by to infer the impact of learning activities to each student's knowledge state.

KT3 contains the required data to model students’ learning behaviour, which could be considered to adjust the difficulty ranking approach. For example, if a student watches a lecture when solving a problem, he will receive less credits than a student who solves the problem individually, even with more attempts and longer time taken.

**References:**

[1] Segal, Avi, et al. "Edurank: A collaborative filtering approach to personalization in e-learning." *Educational Data Mining 2014*. 2014.

[2] Liu, Nathan N., and Qiang Yang. "Eigenrank: a ranking-oriented approach to collaborative filtering." *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. 2008.

[3] A. T. Corbett and J. R. Anderson. Knowledge tracing:  
Modeling the acquisition of procedural knowledge. Usermodeling and user-adapted interaction, 4(4):253–278, 1994.

[4] Segal, Avi, et al. "A difficulty ranking approach to personalization in E-learning." *International Journal of Human-Computer Studies* 130 (2019): 261-272.

[5] David, Yossi Ben, Avi Segal, and Ya'akov Gal. "Sequencing educational content in classrooms using Bayesian knowledge tracing." *Proceedings of the sixth international conference on Learning Analytics & Knowledge*. 2016.

[6] Z. A. Pardos and N. T. Heffernan. KT-IDEM: Introducing Item Difficulty to the Knowledge Tracing Model. In UMAP, 2011.

[7] Segal, Avi, et al. "Combining Difficulty Ranking with Multi-Armed Bandits to Sequence Educational Content." *arXiv e-prints* (2018): arXiv-1804.

[8] Choi, Youngduck, et al. "Ednet: A large-scale hierarchical dataset in education." *International Conference on Artificial Intelligence in Education*. Springer, Cham, 2020.

[9] Schein, Andrew I., et al. "Methods and metrics for cold-start recommendations." *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. 2002.

[10] <https://santatoeic.com/>

[11] <https://www.dropbox.com/s/ph9rnwkjao72k60/RL_LAK_22%20%2814%29.pdf?dl=0>