

Modelling Math Learning on an Open Access Intelligent Tutor

David Azcona¹, I-Han Hsiao², and Alan F. Smeaton¹

¹ Insight Centre for Data Analytics, Dublin City University

² School of Computing, Informatics & Decision Systems Engineering, Arizona State University

Abstract. This paper presents a methodology to analyze large amount of students' learning states on two math courses offered by Global Freshman Academy program at Arizona State University. These two courses utilised ALEKS (Assessment and Learning in Knowledge Spaces) Artificial Intelligence technology to facilitate massive open online learning. We explore social network analysis and unsupervised learning approaches (such as probabilistic graphical models) on these type of Intelligent Tutoring Systems to examine the potential of the embedding representations on students learning.

Keywords: Machine Learning, Intelligent Tutoring Systems, Social Network Analysis, MOOC

1 Introduction

Massive Open Online Courses (MOOCs) are revolutionizing education by giving students around the world open access to first-class education via the web. Lectures, readings, exercises and discussion forums are now one click away to anybody, anywhere. In 2016, Arizona State University (ASU) launched the Global Freshman Academy (GFA) where they provide first-year university courses through the edX platform allowing students to earn transferable ASU credits from anywhere. ASU currently offer 13 courses and our analysis will focus on two Math modules: College Algebra and Problem Solving, and Precalculus. These courses leverage the Assessment and Learning in Knowledge Spaces (ALEKS) technology, which is a web-based artificially intelligent assessment and learning system. ASU's GFA Math courses combined with ALEKS technology is ASU's effort to get students ready for college-level mathematics. The effectiveness and adaptiveness of this Artificial Intelligence (AI) tutoring systems have the potential to motivate and help students acquire these skills.

2 Related work

ALEKS leverages AI to map student's knowledge. ALEKS is based on knowledge spaces, which was introduced in 1985 by Doignon and Falmagne, who describe

the possible states of knowledge of a learner [4]. In order to develop a knowledge space, a domain like Algebra or Chemistry is modelled and divided into a set of concepts and feasible states of knowledge where the student’s knowledge is at any given time. This technology adapts and navigates the students by determining what the student may know and may not know in a course and guides her to the topics she is most ready to learn. It assess the student’s knowledge periodically to ensure topics are learned and retained [5]. Recent research has shown using ALEKS for learning Math has a positive learning impact on an after-school program for more than 200 sixth graders [2, 3].

Recent research has shown that Social Network Analysis (SNA) measurements can be predictive features for machine learning models in addition to generic content-based features [1]. Moreover, sequential modelling (i.e. Hidden Markov Models (HMMs)) can be useful to uncover student progress or students’ learning behaviours [8][7], [6]. We hypothesise that by modelling the evolution of large amount student’s working behaviours with social network features, will allow us to uncover student’s progression and the possibility to enhance student experience with further personalized interventions on these Intelligent Tutoring Systems as it gathers rich information about concepts, topics and learning states.

3 Data collection

An anonymized data of 15,000+ students learning on the two Math courses in edX with the ALEKS technology was collected between April 2016 and October 2017. Students are assessed continuously while navigating through ALEKS and daily aggregates of the topics learned and retained are generated. We tracked 40,000+ assessments and 8+ million daily aggregates. In addition, 5+ million transactions of students navigating through the concepts have been extracted from the edX logs. Each timestamped transaction contains the student, the concept being studied and a learning state. The learning states are the following and final states are determined by the system:

- **L**: Initial state for each concept where a student reads the **L**esson
- **C**: Intermediate state where a student gets an exercise **C**orrect
- **W**: Intermediate state where a student gets an exercise **W**rong or **I**ncorrect
- **E**: Intermediate state where a student asks for a working **E**xample
- **S**: Final state where a student has **M**astered a particular concept
- **F**: Final state where a student has **F**ailed to master a particular concept

4 Methodology

Through the ALEKS platform, we then observe the learning states students go through. We analysed how students transition from one learning state to another and in between concepts and slices (higher-level representations of concepts). The data has been stored and indexed in a non-relational database and made available using a web application where they can explore individual students

and how they are redirected through the material, concepts and their underlying difficulty based on completion and likely sequence patterns around concepts and slices. Figure 1 shows the time distribution for each learning state for a particular student with respect to the other states.

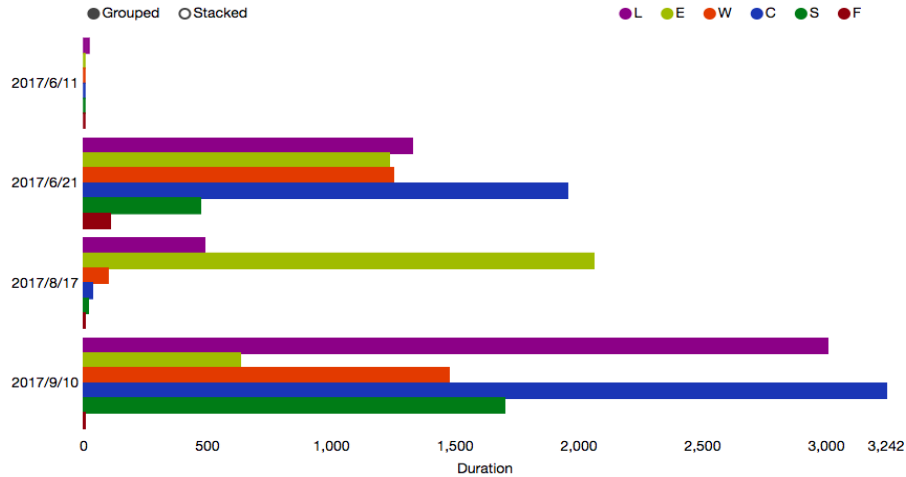


Fig. 1. Grouped transactions for a particular student

4.1 Network Analysis

We follow a social network analysis approach to investigate the structure of the student's learning on these systems by leveraging networks and graph theory. Network analysis can give us insights into the properties of these learning states. A network of the observed learning states is developed for each student and each topic and posted on the web application. Figure 2 shows two very different learning paths from students working on similar topics, a linear learning path versus a student who failed to master the concept in the beginning but was brought back to the concept and mastered it. Considering these learning states as nodes of a network enables us to analyse their centrality and measure their importance in the learning of these concepts and disciplines. By creating a network for each topic with all student data, we are measuring the degree centrality of the learning states and finding out where students are most likely to struggle or succeed by looking at the final states. Moreover, we are developing a network with all the learning states and their corresponding topics and analysing the students' progression; looking for centrality hubs; hypothesizing whether further interventions may be added; measuring eigenvector centrality and how important are the neighbors of the nodes; using the geodesic distance to measure the closeness and betweenness between topics for each slice, which slices take more effort from

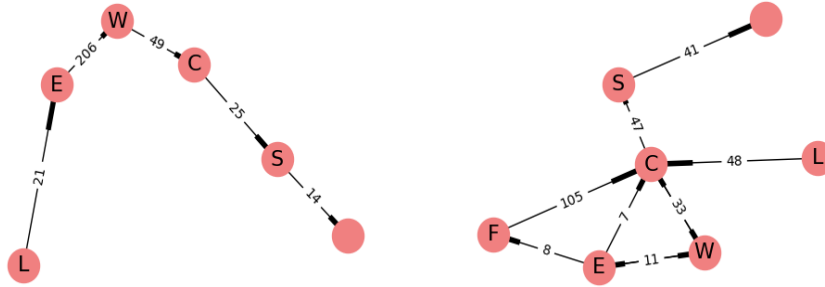


Fig. 2. Networks of learning states where students are navigated

students and how to help them succeed. In addition, we will cluster the nodes and see if the groupings correspond with the sections of the topics.

4.2 Summary & Future work

The learning states extracted from the transactions and the underlying navigation through the course can be modelled using a Markovian procedure by assuming the future learning states depend only on the current learning state. This could be further developed by looking at sequences of learning states and topics learned in order to model the likelihood of future learning states. In our case we consider the learning states as observable states and are modelling their learning using an HMM, the unobserved or hidden states are estimated from the sequence of learning states students follow and are navigated on the system. On the other hand, based on the large number of transactions gathered, another higher-level representation of this sequence of learning states can be learned using embeddings. We are learning these embeddings by inputting the sequence of learning states to a supervised machine learning algorithm, a Recurrent Neural Network, that take the time constraint into consideration. This is a promising avenue of research that has successfully been applied to Computer Vision to learn high-level features using Deep Learning approaches.

Acknowledgements: This research was supported by the Irish Research Council in association with the National Forum for the Enhancement of Teaching and Learning in Ireland under project number GOIPG/2015/3497, by Science Foundation Ireland under grant SFI/12/RC/2289, and by Fulbright Ireland. The authors are indebted to the Action Lab at EdPlus in Arizona State University for their help.

References

1. Chen, W., Brinton, C.G., Cao, D., Mason-singh, A., Lu, C., Chiang, M.: Early detection prediction of learning outcomes in online short-courses via learning behaviors. *IEEE Transactions on Learning Technologies* (2018)
2. Craig, S.D., Anderson, C., Bargaglioitti, A., Graesser, A.C., Okwumabua, T., Sterbinsky, A., Hu, X.: Learning with aleks: The impact of students' attendance in a mathematics after-school program. In: *International Conference on Artificial Intelligence in Education*. pp. 435–437. Springer (2011)
3. Craig, S.D., Hu, X., Graesser, A.C., Bargagliotti, A.E., Sterbinsky, A., Cheney, K.R., Okwumabua, T.: The impact of a technology-based mathematics after-school program using aleks on student's knowledge and behaviors. *Computers & Education* **68**, 495–504 (2013)
4. Doignon, J.P., Falmagne, J.C.: Spaces for the assessment of knowledge. *International journal of man-machine studies* **23**(2), 175–196 (1985)
5. Doignon, J.P., Falmagne, J.C., Cosyn, E.: Learning spaces: A mathematical compendium. In: *Knowledge Spaces*, pp. 131–145. Springer (2013)
6. Hsiao, I.H., Huang, P.K., Murphy, H.: Uncovering reviewing and reflecting behaviors from paper-based formal assessment. In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. pp. 319–328. ACM (2017)
7. Piech, C., Huang, J., Nguyen, A., Phulsuksombati, M., Sahami, M., Guibas, L.: Learning program embeddings to propagate feedback on student code. arXiv preprint arXiv:1505.05969 (2015)
8. Piech, C., Sahami, M., Koller, D., Cooper, S., Blikstein, P.: Modeling how students learn to program. In: *Proceedings of the 43rd ACM technical symposium on Computer Science Education*. pp. 153–160. ACM (2012)