Learning to Teach like a Bandit

Mykola Pechenizkiy
Department of Computer Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven
The Netherlands
m.pechenizkiy@tue.nl

Pedro A. Toledo
Department of Computer Science
Universidad de La Laguna
Fac. Física. Francisco Sánchez SN
Santa Cruz de Tenerife, Spain
pedro@isaatc.ull.es

ABSTRACT
Designing a good course curriculum is a non-trivial task many teachers have to deal with on a regular basis. There are multiple learning methodologies available, but some of the basics are common; thus, one of the important steps is to identify key concepts and knowledge or skills prerequisites for mastering them. If this can be done properly, a teacher acting as a course designer can think how to sequence the material. After the first edition of the course the teacher takes into account what went well and what adjustments to the course curriculum would be appropriate. With the growing popularity of ITS and recently MOOCs there are more opportunities for data-driven decisions on how to sequence learning materials and activities to optimize the learning process. Personalizing curriculum to different students is also becoming possible based on how well students learn or are expected to learn. Finding the best possible curriculum for all, a group or an individual student is a nontrivial problem that has an explore-exploit nature. We can use ideas of reinforcement learning and consider course design and learning activities sequencing as a kind of multi-armed bandit problem. We illustrate how to sequence these activities iteratively by employing the genetic process mining framework for generating a population of curriculum candidates from historical data and how to choose these candidates using the Bandit strategy to address the exploration-exploitation trade-off.

Keywords

Iterative course design, reinforcement learning, process mining.

1. INTRODUCTION
The design of a course curriculum or more broadly instructional design has been traditionally an important challenge to educators or teachers responsible for construction of a course [2,3]. And it is a key part of getting satisfactory results in any teaching and learning process. Therefore, teachers have to invest significant time and effort in it.

The design process usually takes a long period of time and it is done using what is called a **curriculum development and implementation cycle** [1] as illustrated in Figure 1.

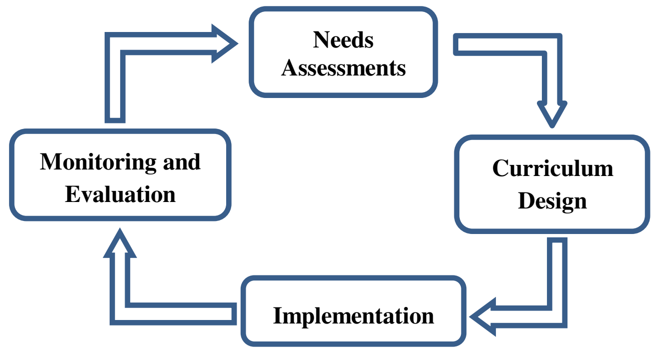

Figure 1 Course Curriculum Design Cycle

The design process may result with a high number of course curriculum alternatives with different task sequencing [4, 5], which would be difficult to evaluate and compare to each other accurately. We propose a new strategy for the curriculum design to construct a data-driven process [8] which may automatically justify changes in the curriculum design.

2. LEARNING TO TEACH LIKE A BANDIT APPROACH
Our course curriculum model describes the relations among the set of activities that may take place during a course, and the resources used with them. To build it, we follow the classical curriculum design cycle. In the first phase (Needs Assessments), a set of curriculum designs are originally proposed by the course teachers based on the course assessments. In the second phase (Curriculum Design), the population is considered the starting population of solutions for the Genetic Process Mining Algorithm. Genetic Process Mining is the technique designed by applying genetic algorithms to perform process mining. In the third phase (Implementation), once new evolved curricula have been obtained, one of them is selected using the Bandit Algorithm. The Multi-Armed Bandit Problem can be modeled as a Single-State Markov Decision Process [6, 7]. To choose, it considers the expected performance distribution of each of the curriculums from the population. The selected curriculum is implemented, and new information is gathered. Finally in the fourth phase (Monitoring and Evaluation), the new data is used to update the curriculum performance information. Back to the first stage of the cycle, the curriculum population is pruned using the updated expected performance distribution. A graphical representation of the approach can be seen in Figure 2.

As a consequence of the running cycle, a multi-objective optimization of the model takes place. Each curriculum selection stage has the consequence of gathering new data from the high performance curriculum implementations. In the long term, the obtained data log will mainly correspond to high performance curriculum implementations. Additionally, each time the Genetic Process Mining algorithm is executed, the obtained models are optimized to fit the gathered data. Consequently, the populations of models obtained in the long-term execution of the cycle correspond to high performance curriculum models that accurately describe the real process implemented for the course.

The described method present serious advantages with respect to the straightforward process mining approach:

Firstly, the inclusion of Multi-Armed Bandit Problem Strategy solves the curriculum selection problem in the exploration exploitation scenario.

Secondly, a continuous strategy of curriculum improvement is defined. This is so because the evolutionary algorithm will use the gathered performance information to transform the population of curriculum models.

Finally, the combination of the Multi-Armed Bandit Problem Strategy and Genetic Process Mining solves the multi-objective problem of curriculum mining. This problem consists on maximizing the model fitting to the data, and the expected performance.

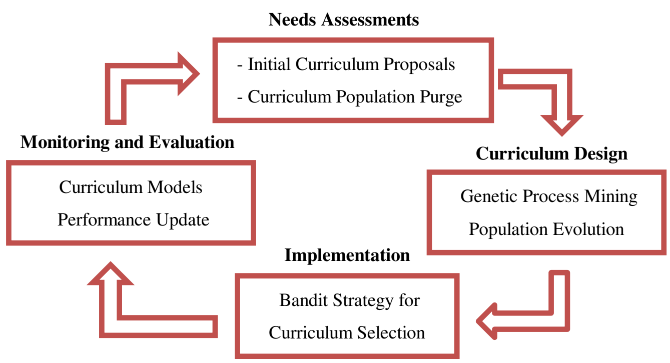

Figure 2 Learning to Teach as a Bandit Approach

2.1 Application Settings

Different scenarios are suitable for the application of the 'Learning to Tech as a Bandit' strategy.

On the one hand, when a new course is designed, multiple possibilities for the curriculum definition are considered. In this case the successive iterations of the course implementation could be used to improve the curriculum design. At each iteration, the whole classroom would follow the selected curriculum until the end. Only after the course is finished, the obtained data from the students' results is analyzed. The information about the evaluation of the implemented curriculum is updated and, based on it, the next curriculum for the following course is implemented. Obviously, it is possible to plan simultaneous courses implementing each one a different curriculum. In that case the curriculum evolution is faster in terms of time spent to discover better solutions. However, higher costs in terms of implementation of poor performance curriculum are also expected.

On the other hand, the scenario of student curriculum personalization in MOOCs is also a plausible setting for the approach. In this case the purpose is not to get the better curriculum design for the course. Furthermore, it is assumed that there is not a singular curriculum to use in a one-fits-all mode. Usually, the group of students of a MOOC is heterogeneous, and therefore the use of personalized curriculum for each user is advisable. In this case, the strategy would consist on obtaining the best fitting curriculum for each.

3. CONCLUSIONS AND FUTURE WORK

This paper proposes a novel approach for curriculum design and adaptation. The hard problem of curriculum design has been formalized in terms of Curriculum Mining. Furthermore, the problem of searching the best design has been stated, pointing out the exploration-exploitation trade off. The approach tackles this issue, with the formulation of a Multi-Armed Bandit Problem. Consequently, the cost of implementation of suboptimal curriculum is minimized. The approach is valid for different settings. It may be either used in the context of new course curriculum design or for student curriculum personalization.

The proposal would need empirical validation to contrast its efficiency in comparison with the traditional curriculum design and personalization methods. Nevertheless, all the ingredients necessary for the implementation are already available: not only the Genetic Process Mining algorithms but also the Multi-Armed Bandit Problem solvers. Therefore, implementing the approach results a straightforward task.

Learning to Teach as a Bandit, is a first attempt to build a data-driven approach to curriculum design, incorporating the scientific method to a traditionally experience based task. The whole educational community would benefit from the approach advantages. Teachers would have a solid guide for curriculum design and a clear method to improve their courses. Additionally, students could have not only courses with a further adaptation to their expectations, but also they could minimize the experience of poorly designed curricula.

4. REFERENCES

[1] Frase, L.E., English, F. W., Poston Jr., W. K. (2000) The Curriculum Management Audit. Improving School Quality. Rowman & Littlefield Education.

[2] Gagné, R. M., & Driscoll, M. P. (1988). Essentials of learning for instruction. Englewood Cliffs, NJ: Prentice-Hall

[3] Gagne, R., Briggs, L. & Wager, W. (1992). Principles of instructional design (4th Ed.). Fort Worth, TX: HBJ College.

[4] Morales, R., Agüera, S. (2002) Dynamic Sequencing of Learning Objects. IEEE International Conference of Advanced Learning Technologies (ICALT)

[5] Wiley, D.A. (2000) Learning Object Design and Sequencing Theory. PhD Dissertation. Birgham Young University.

[6] Mitchell, C.M., Boyer, K. E., Lester, J. C. (2013) A Markov Decision Process Model of Tutorial Intervention in Task-Oriented Dialogue. Artificial Intelligence in Education. LNCS. Vol. 7926, pp 828-831.

[7] Almond, R.G. (2007) An Illustration of the Use of Markov Decision Processes to Represent Student Growth (Learning). Reseach Report. ETS, Princeton, NJ

[8] Koedinger, K. R., Brunskill, E., Baker, R., McLaughlin, E. A., Stamper, J. (2013). New Potentials for Data-Driven Intelligent Tutoring System Development and Optimization. AI Magazine. Vol 34, No 3,3, 27-41.