Online Optimization and Personalization of Teaching Sequences

Benjamin Clement, Didier Roy, Manuel Lopes, Pierre-Yves Oudeyer

Online Optimization and Personalization of Teaching Sequences

Benjamin Clément[1], Didier Roy[1], Manuel Lopes[1], Pierre-Yves Oudeyer[1]

[1] Flowers Lab – Inria Bordeaux Sud-Ouest, Bordeaux 33400, France,

didier.roy@inria.fr

1 INTRODUCTION

Intelligent Tutoring Systems (ITS) have been proposed to make education more accessible, more effective and simultaneously as a way to provide useful objective metrics on learning. More and more projects started on Massive Open Online Course (MOOC) and learning applications for mobile devices.

Optimized ITS are great assets for training and can go further and participate effectively in the fight against school dropout, provided that the content is personalized for each learner, and intrinsec motivation is stimulated.

Traditional School:

- 1 teacher
- >30 different students

For everyone:

- The same activities
- In the same sequence
- At the same speed

School 3.0:

- 1 teacher
- > 30 different students
- > 30 digital system to help

For each:

- Personalized activities
- In the sequence that help them improve faster
- Considering the motivation and frustation of the student

➔ personalized education, fight school dropout

According to [8], there are four main components of an ITS: i) a *cognitive model* that defines the domain knowledge or which steps need to be made to solve problems in a particular domain; ii) a *student model* that considers how students learn, what is the evolution of their cognitive state depending on teaching activities; iii) a *tutoring model* that defines, based on the cognitive and the student models, what teaching activities to present to students ; iv) a user interface model that represents how the interaction with the students occurs, how problems are proposed to the learners.

In this work we are more focused on the *tutoring model*, that is, how to choose the activities that provide a better learning experience based on the estimation of the student competence levels and progression, and some knowledge about the cognitive and student model. We can imagine a student wanting to acquire many different skills, e.g. adding, subtracting and multiplying numbers. A teacher can help by proposing activities such as: multiple choice questions, abstract operations to compute with a pencil, games where items need to be counted through manipulation, videos, or others. The challenge is to decide what is the optimal sequence of activities that maximizes the average competence level over all skills.

There are several approaches to develop a *Tutoring Model*. A first approach is based on hand-made optimization and on pedagogical theory, experience and domain knowledge. There are many works on this [4], [8]. A second approach, more relevant for our work, is that the optimization is made automatically without particular assumptions about the students or the knowledge domain.

Our ITS system aims at providing to each student the activities that are giving the highest learning progress. Learning progress provided by activities is estimated at runtime and use the students results. This approach has three main advantages.

No need for a precise cognitive/student model. In most cases the tutoring model incorporates a student model inside. Given students' particularities, it is often highly difficult or impossible for a teacher to understand all the difficulties and strengths of individual students and thus predict which activities provide them with maximal learning progress. Also, identifying all such parameters for a single student is a very hard problem due to the lack of data, the intractability of the problem, and the lack of identifiability of many parameters. The diversity in learning styles, attitudes, and cultures brings additional difficulties to define student models that are relevant for large scale deployement. This often results in models which are inaccurate in practice [2]. Also, it has been shown that a sequence that is optimal for the average student is often suboptimal for most students, from the least to the most skilled [6].

We consider that it is important to be as independent as possible of a cognitive and student model. This requires that the ITS explores and experiments various activities to estimate their potential for learning progress for each student.

Efficient optimization methods. We present methods that do not make specific assumptions about how students learn and only require information about the estimated learning progress of each activity. We make a simple assumption that activities that are currently estimated to provide a good learning gain, must be selected more often. A very efficient and well studied formalism for these kind of problems is Multi-Armed Bandits [3]. Following a casino analogy, at each step we can choose a slot-machine and we get to observe the payback we get, the goal it to find the best arm, but while we are trying to discover it we have to bet to test them.

More motivating experience for the students. Our approach considers that exercises which are currently providing the higher learning progress must be the ones proposed. This allows not only to use more efficient optimization algorithms but also to provide a more motivating experience to students. Several works in psychology and neuroscience have argued that the human brain feels intrinsic pleasure in practicing activities of optimal difficulty, i.e. neither too easy nor too difficult, but slightly beyond the current abilities, also known as the zone of proximal development [9,10].

Our main contributions are the use of highly performing Multi-Armed Bandit algorithms, a simpler factored representation of the cognitive model that maps activities to the minimum necessary competence levels, and considering that the acquisition of a Knowledge Component (KC) is not a binary variable but defined as the level of comprehension of that KC. The advantage of using Multi-Armed Bandit (MAB) algorithms is that they are computational efficient and require a weaker dependency between the tutoring and the cognitive and student models.

2 ITS WITH MULTI-ARMED BANDITS

Relation between KC and pedagogical activities. In general, activities may differ along several dimensions and may take several forms. Each activity can provide opportunities to acquire different skills/knowledge units, and may contribute differentially to improvement over several KCs. While certain regularities of this relation may exist across individuals, it will differ in detail for every student. Still, an ITS might use this relation in order to estimate the level of each student. We will latter show how to further simplify this assumption.

Estimating the impact of activities over students' competence level in knowledge units. Key to the approach is the estimation of the impact of each activity over the student's competence level in each knowledge unit. This requires an estimation of the current competence level of the student for each KCi. We do not want to introduce, outside activities, regular tests that would be specific to evaluate each KCi since it would have a high probability to negatively interfere with the learning experience of the student. Thus, competence levels need to be inferred through stealth assessment that uses indirect information coming from the combination of performances in activities. Indeed, we assume here that this is a good indicator of the learning progress over KCi resulting from doing an activity with parameters. If you have repeated successes in an activity for which the required competence level is higher than your current estimated competence level, this means you are probably progressing.

2.1 RiARiT Algorithm: Right Activity at Right Time

We here use and adapt MAB approaches to the problem of optimal teaching, where the gambler is replaced by the teacher, the choice of machines is replaced by a choice of learning activity, and money is replaced by learning progress (which is a proxy for maximizing acquired skills). A particularity here is that the reward (learning progress) is non-stationary, which requires specific mechanisms to track its evolution. Indeed, here a given exercise will stop providing reward, or learning progress, after the student reaches a certain competence level. Also we cannot assume that the rewards are i.i.d. as different students will have different preferences and many human factors, i.e. distraction, mistakes on using the system, create several spurious effects. Thus, we rely here on a variant of the algorithm (EXP4 [1], [3]).

2.2 ZPDES: Zone of Proximal Development and Empirical Success

Our goal is to reduce the dependency on the cognitive and student models and so we will try to simplify further the algorithm. Our simplification will use the concept of zone of proximal development and the empirical estimation of learning progress.

As discussed before focusing teaching in activities that are providing more learning progress can act as a strong motivational cue. We use too the concept of the zone of proximal development [5] that considers that activities that are slightly beyond the current abilities of the learner are the more motivating.

Estimating explicitly how the success rate on each exercise is improving will remove the dependency on the links between activities parameters and skills levels of the students. We compare more recent success with all the previous past, providing an empirical measure of how the success rate is increasing. We no longer estimate the competence level of the student, and directly use the reward estimation.

This concept will provide three advantages: improve motivation; further reduce the need of quantitative measures for the educational design expert; and provide sequence of activities that follow a more sequential order.

3 Simulations and Results

We use a specific teaching scenario about learning how to use money, typically targeted to students of 7-8 years old, more detailed in [7].
We present a set of simulations with virtual students to test systematically different properties of our algorithm. The results show how fast and efficiently our algorithms estimate and propose exercises at the correct level of the students. Each experiment considers a population of 1000 students generated using the previous methods and lets each student solve 100 exercises. RiARiT and ZPDES are both better than Predefined. This is explained because when the student is not able to understand exercises with a specific parameter, a predefined sequence can not adapt and propose an alternative path. The figure 1 shows the skill's levels evolution during 100 steps.

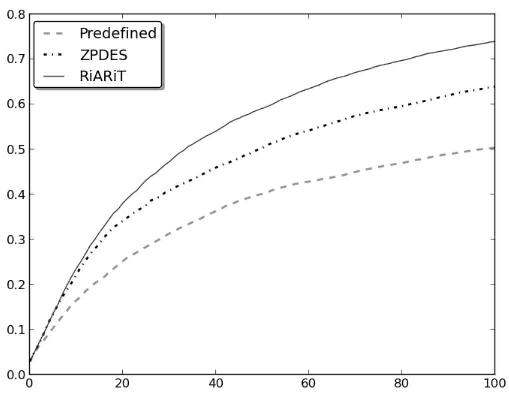

Fig. 1. RiARiT et ZPDES are better than predefined sequence: the ITS is optimized.

4 Conclusion and future work

With the use of bandit algorithms to select learning activities which empirically maximize student's learning progress, we proposed a new approach to intelligent tutoring systems.

We showed through simulations and empirical results that a very efficient algorithm, that tracks the learning progress of students and proposes exercises proportionally to the learning progress, can achieve very good results.

In order to evaluate our algorithms, we use as baseline an optimized sequence created based on instructional design theory, whose reliability has been validated through several user studies.

After a first experiment with about 100 students in primary schools [7], we are testing deeper this approach in primary schools with about 400 students. As the first experiment, the second one proposes a learning sequence on the money usability. The aim is to compare the performances of the different strategies. Also, two tests, one before and the other after the learning sequence, are evaluating to know if the students make real progresses with these strategies.

5 References

1. P. Auer, N. Cesa-Bianchi, Y. Freund, and R. Schapire. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1):48–77, 2003.
2. J. E. Beck and X. Xiong. Limits to accuracy: How well can we do at student modeling? In Educational Data Mining, 2013.
3. S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and non-stochastic multi-armed bandit problems. *Foundations and Trends in Stochastic Systems*, 1(4), 2012.
4. K.R. Koedinger, E. Brunskill, R. SJd Baker, E. A. McLaughlin, and J. Stamper. New potentials for data-driven intelligent tutoring system development and optimization. AI Magazine, 2013.
5. C. D. Lee. Signifying in the zone of proximal development. *An introduction to Vygotsky*, 2:253–284, 2005.
6. J. Lee and E. Brunskill. The impact on individualizing student models on necessary practice opportunities. In Inter. Conf. on Educational Data Mining, 2012.
7. M. Lopes, B. Clement, D. Roy, and P.-Y. Oudeyer. Multi-armed bandits for intelligent tutoring systems. arXiv:1310.3174 [cs.AI], 2013.
8. R. Nkambou, R. Mizoguchi, and J. Bourdeau. *Advances in intelligent tutoring systems*, volume 308. Springer, 2010.
9. J. Gottlieb, P.-Y. Oudeyer, M. Lopes, A. Baranes. Information Seeking, Curiosity and Attention: Computational and Neural Mechanisms, 2013. *Trends in Cognitive Science*, 17(11), pp. 585-596.
10. E. Deci and R. Ryan, Intrinsic Motivation and Self-Determination in Human Behavior. New York: Plenum, 1985.