

# Multi-Armed Bandits for Intelligent Tutoring Systems

Benjamin Clement  
Inria, France  
benjamin.clement@inria.fr

Didier Roy  
Inria, France  
didier.roy@inria.fr

Pierre-Yves Oudeyer  
Inria, France  
pierre-yves.oudeyer@inria.fr

Manuel Lopes  
Inria, France  
manuel.lopes@inria.fr

## ABSTRACT

We present an approach to Intelligent Tutoring Systems which adaptively personalizes sequences of learning activities to maximize skills acquired by students, taking into account the limited time and motivational resources. At a given point in time, the system proposes to the students the activity which makes them progress faster. We introduce two algorithms that rely on the empirical estimation of the learning progress, RiARiT that uses information about the difficulty of each exercise and ZPDES that uses much less knowledge about the problem.

The system is based on the combination of three approaches. First, it leverages recent models of intrinsically motivated learning by transposing them to active teaching, relying on empirical estimation of learning progress provided by specific activities to particular students. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. The system is evaluated in a scenario where 7–8 year old schoolchildren learn how to decompose numbers while manipulating money. Systematic experiments are presented with simulated students, followed by results of a user study across a population of 400 school children.