Research statement: Inference of human-computation algorithms from massive-scale educational interventions

PhD Student: Matías López y Rosenfeld, e-mail: mlopez@dc.uba.ar

The main goal of the present research statement is to develop an educational computerized framework able to detect in which tasks a child has difficulties and generate a personalized intervention based on automatic observation and evaluation of data, as part of an interdisciplinary project at the crossroads of Computer Science, Cognitive Science, Biology and Psychology.

In Argentina, different government programs have provided most students with a personal laptop computer.This unified digital platform allows educational interventions and research in a country-wide fashion. Implementing this intervention requires the development of many tools and methods. One such method constitutes the main aim of this proposal: the development of data mining techniques to infer human-computation algorithms from the huge corpus of data that is being currently collected.

We aim to combine ideas from intelligent tutor systems developed at Carnegie Mellon University[1] with training cognitive bricks (such as executive functions) at initial levels in primary school using specific games. We will focus on children in their first years of schooling (5-8 years old). By the creation of training games and their deployment country-wide, we propose the formation of a massive-scale repository of human-development cognitive data, and the methods for their analysis.

Specifically, the goals of the proposed work are:

- to detect functions of human computation algorithms during the development at school, such as numerosity, planning, inhibitory control and working memory.
- to design activities to stimulate and identify different human computational capacities and to collect information about how the computations are made while students use these activities
- data mining of the behavioral logs to quantify the involved processes while children learn, looking for unknown new facts about human computation
- the design of new activities and improvement of existing ones to automatically adapt to the child's needs and improve the development of human computation functions.

These objectives build a closed loop in the educational process where the own activities and their usage let us discover how human make computation, which in turn open a path to the development of new activities that will take us a step closer in the path of discovering the secrets of the children's mind architecture.

To achieve the objectives we split the activities in 4 different stages:

1. Identification of cognitive capabilities and generation of training activities.[2,3,4]

2. Desing, development and deployment of a scalable framework, containing the training activities, for interventions and collecting behavioral logs.
3. Mining of collected data: search for correlations in time, and between different activities by mining in massive datasets[5], and develop heuristics to estimate measures and model parameters[6,7].
4. Design adaptive activities: development of machine learning techniques that model individual-user behavior while using the activities. Including adaptive strategies based on ITS[1] that will let us calibrate difficulty and other parameters automatically within the activities. The main goal will be the detection of students that try to solve problems without understanding their nature[8].

Nowadays stage 1 has been finished and stage 2 will be finished by the end of 2014.

References

1. Koedinger, K., Corbett, A.: Cognitive tutors: Technology bringing learning science to the classroom. The Cambridge handbook of the learning sciences (2006) 61–78
2. Wilson, A., Revkin, S., Cohen, D., Cohen, L., Dehaene, S., et al.: An open trial assessment of "the number race," an adaptive computer game for remediation of dyscalculia. Behavioral and Brain Functions **2**(1) (2006) 20
3. McCandliss, B., Beck, I., Sandak, R., Perfetti, C.: Focusing attention on decoding for children with poor reading skills: Design and preliminary tests of the word building intervention. Scientific Studies of Reading **7**(1) (2003) 75–104
4. Goldin, A., Pezzatti, L., Battro, A., Sigman, M.: From ancient greece to modern education: Universality and lack of generalization of the socratic dialogue. Mind, Brain, and Education **5**(4) (2011) 180–185
5. Rajaraman, A., Ullman, J.: Mining of massive datasets. Cambridge Univ Pr (2011)
6. Andoni, A., Indyk, P.: Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In: Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on, Ieee (2006) 459–468
7. Charikar, M.S.: Similarity estimation techniques from rounding algorithms. In: Proceedings of the thiry-fourth annual ACM symposium on Theory of computing, ACM (2002) 380–388
8. Baker, R., Corbett, A., Koedinger, K., Evenson, S., Roll, I., Wagner, A., Naim, M., Raspat, J., Baker, D., Beck, J.: Adapting to when students game an intelligent tutoring system. In: Intelligent Tutoring Systems, Springer (2006) 392–401