Adaptive systems for real-life education need explicit domain and activity models – and ways to generate them automatically

Detmar Meurers, Björn Rudzewitz, Martí Quixal

Digital systems that are capable of adaptively guiding learners hold substantial promise for speeding up learning and improving learning outcomes for individuals in heterogeneous target groups. Depending on the characteristics of a given learner, the idea behind adaptive Intelligent Tutoring Systems (ITS) is a) to select the next learning activity so that it is appropriately challenging (macro adaptivity), and b) to support the learner's work on a given activity with scaffolded feedback (micro adaptivity). While the first generation of ITS was successful in effectively fostering learning (Kulik & Fletcher, 2016), that success has been limited to specific subject domains. For each subject domain to be learned, an explicit domain model needs to be created (Sottilare et al., 2016) as basis for learner and activity models. So the creation of a new ITS generally involves a very substantial manual effort in an extensive cross-disciplinary collaboration with domain and pedagogical experts. As a result, even though ITS have been around for over 30 years, very few systems have been developed for real life use, and their development has largely been limited to so-called well-defined domains (e.g., algebra, geometry, programming, physics) which lend themselves to such explicit domain modeling.

More recently, companies such as Area 9 Lyceum have answered the call by the German KMK to develop adaptive systems for the German school context and propose to provide a platform that is independent of the content domain ("Eine schülerzentrierte Lernplattform für alle Alterngruppen, Lehrpläne und Interessenvertreter [A student-centered learning platform for all age groups, curricula, and stakeholders]", http://area9lyceum.de). The underlying approach is to determine the difficulty of a learning step by observing students as they complete activities, successfully or not, and to ask them about their perception of difficulty. Such an approach essentially follows that of Computer Adaptive Testing (CAT), which speeds up testing by adaptively selecting those test items that discriminate best at the level of a given person taking the test. However, the standard models underlying such CAT systems crucially assume test unidimensionality, i.e., there must be one dominant factor that accounts for item performance (Hambleton et al., 1991).

In this talk, we argue that for an adaptive learning environment such an assumption is fundamentally flawed and that explicit, rich domain models as found in the original ITS designs are crucial for effective micro- and macro-adaptivity. We start by presenting a new learning analytic study based on the log data collected during the FeedBook study (Meurers et al., 2019), a full year randomized controlled field study with 255 students from 10 seventh-grade English classes in four German high schools. We identify four behavioral patterns and differential effects of specific feedback on ultimate achievement related to these learner clusters. While this confirms the value of taking such behavioral patterns into account, we then turn to exploring the range of factors that need to be taken into account in an activity model capable of intrinsically characterizing activity complexity in a way that supports macro-adaptive activity selection. Different from the first generation ITS systems, Natural Language Processing techniques make it possible to automatically derive the rich activity models needed to link an activity to the pedagogical learning targets spelled out in the curriculum (Quixal et al., 2021). We end with an outlook on the macro-adaptive activity sequencing algorithm currently being evaluated in a study and our vision for automatically generating different parametrizations of activities automatically to provide the high number of activities one needs for an adaptive system capable of matching the needs of a heterogeneous learner population.

Literatur

Hambleton, R. K., J. N. Zaal & J. P. Pieters (1991). Computerized adaptive testing: Theory, applications, and standards. In *Advances in educational and psychological testing: Theory and applications*, Springer, pp. 341–366.

Kulik, J. A. & J. Fletcher (2016). Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of Educational Research* 86(1), 42–78.

Meurers, D., K. De Kuthy, F. Nuxoll, B. Rudzewitz & R. Ziai (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics* 39, 161–188. URL https://doi.org/10.1017/S0267190519000126.

Quixal, M., B. Rudzewitz, E. Bear & D. Meurers (2021). Automatic annotation of curricular language targets to enrich activity models and support both pedagogy and adaptive systems. In *Proceedings of the 10th Workshop on NLP for Computer Assisted Language Learning*. pp. 15–27. URL https://aclanthology.org/2021.nlp4call-1.2.pdf.

Sottilare, R. A., A. C. Graesser, X. Hu, A. Olney, B. Nye & A. M. Sinatra (2016). *Design recommendations for intelligent tutoring systems: Volume 4-domain modeling*, vol. 4. US Army Research Laboratory.