Standardizing Modeling of User Behaviors: Which Behaviors Matter?

Ryan S. Baker[1], Chad J. Coleman[2]
University of Pennsylvania[1], Teachers College Columbia University[2]
pTopical Area: Learner Modeling

INTRODUCTION

2004 marked the publication of what were arguably the first automated models of disengaged behavior within an adaptive instructional system (AIS) – models that captured gaming the system (Baker, Corbett, & Koedinger, 2004), or a subcategory of gaming, hint abuse (Aleven, McLaren, Roll, & Koedinger, 2006). Since then there has been a great deal of work that aimed to measure a range of disengaged behaviors (see review in Baker & Rossi, 2013), as well as behaviors that reflect positive self-regulation during learning (Shih, Koedinger, & Scheines, 2011; Tscholl et al., 2016). This work has shown not only that these behaviors can be modelled, but also that the resultant models can predict differences in learners' outcomes, both in the short-term (Rowe, McQuiggan, Robison, 2009; Sabourin, J., Rowe, Mott, & Lester, 2011) and in the long-term (San Pedro et al., 2013; Pardos et al.. 2014). Finally, many of these behaviors are amenable to intervention to reduce disengagement and increase the frequency of positive behaviors (e.g. Baker et al, 2006; Arroyo et al., 2007; Bouchet et al., 2016), creating the potential to improve student outcomes.

PROBLEM OR OPPORTUNITY FOR STANDARDIZATION

Despite all of this research, the set of disengaged and engaged behaviors being modelled varies system-to-system, and the choice of behaviors modelled in specific research projects remains ad-hoc and unsystematic. There have been no efforts that the authors are aware of to standardize what set of behaviors should be modelled. This is partly because the field is still determining which behaviors matter, and new behaviors of importance are likely to emerge, especially as the design of adaptive instructional systems changes to accommodate technologies such as virtual and augmented reality, and as learners themselves change over time. Nonetheless, many behaviors have been identified and studied in a range of learning systems – for example, models of gaming the system have been developed in at least 7 AIS systems that the authors are aware of, and models of off-task behaviour have been developed in at least 9 AIS systems that the authors are aware of.

Given all of the work to detect specific known behaviors such as these, it seems feasible to start moving as a community to setting up a taxonomy of behaviors that should generally be modelled. Doing so, and having a shared framework for disengaged and engaged behaviour in adaptive instructional systems, will enable different instructional systems to communicate with each other in a consistent way about how often students engage in specific behaviors known to be important, such as – for instance – gaming the system.

In this presentation, we will summarize the literature on these behaviors; which can be modeled; which have evidence for their importance to learner outcomes; and how robust and general that evidence is. We will also present thoughts towards a framework for modeling behavior in adaptive instructional systems, which would encapsulate widely-studied behaviors and have sufficient flexibility to accommodate new discoveries about user behaviors that emerge as adaptive instructional systems change, and the learners who use them change as well.

To start, this framework may be as simple as a list of common behaviors and their manifestations, with a commitment to representing behaviors using this taxonomy when possible, and embedding the taxonomy in a standard. This standard can and should be designed in a fashion where new behaviors can be tagged and added easily within the system, either for a specific system or as an addition to the standard by general agreement across a broader community of stakeholders.

The framework could support the communication of information between AIS systems with reference to this taxonomy, perhaps as a probability or estimated proportion of time that a student engaged in each behavior of interest (or an indicator if the behavior is inapplicable within a specific system). If a framework were designed in this fashion, it could accommodate outputs from a range of models in an algorithm/approach-neutral fashion. Whether or not a model of a specific behaviour was based on feature-engineered machine learning, deep learning, rational modelling/knowledge engineering, or some as-yet-unknown approach, it could provide probability estimates or estimated proportions of time that would be meaningful and useful to the next AIS system the learner would encounter. The models themselves would not need to be domain or system-general; true system-generality is as yet uncommon in this space, with only Paquette et al.'s (2015) gaming detector being validated to be general in this fashion. Paquette et al's detector itself has only been validated to work effectively for three AIS systems (Paquette et al., 2015). Having a framework like this would be a step towards integrating information about a learner across multiple AIS systems, towards getting a broader picture of the learner, including their propensities to disengage in certain types of learning activities (perhaps some learners disengage with didactic materials whereas others disengage with games), their propensities to disengage more for some topics than others (an indirect measure of interest), and their overall path towards long-term engagement or disengagement with the learning subject.

DISCUSSION

One of the largest arguments against creating a cross-system framework for disengaged and engaged behavior is that new behaviors may still be discovered in the years to come, particularly as the systems we study change. However, we would argue that several behaviors have been shown to matter in several contexts, and are worth developing standards for. Furthermore, it is not clear when or even ***if*** new behaviors will stop being discovered; the designs of AIS systems may continue to change for decades as new interaction technologies are developed, and changes in design may drive ongoing changes in how learners engage and disengage. This argues for a flexible and expansible framework rather than rejecting an attempt to codify and build on what is today known.

A second argument is that a cross-system framework will not be used, because different systems will be unable to trust each others' models, and will not want to use evidence from other systems. It is true that cross-AIS transfer of knowledge about students has not yet become a practice. However, many systems have models of these behaviors that are published and heavily validated. Furthermore, the long-term predictive power of these behaviors (e.g. San Pedro et al., 2013; Pardos et al., 2014) suggests that a student who is disengaged in 8^{th}-grade mathematics is likely to remain disengaged in 9^{th}-grade. So, too, a military trainee's disengagement is relatively unlikely to simply disappear, from one month to the next. As such, retaining information on a learner's previous-year disengagement is likely to give a newly-encountered AIS system a head start on adapting to that learner's needs.

The biggest benefit to creating a cross-system framework and standard for engaged and disengaged behavior is the potential, discussed above, for AIS systems to exchange information and build information over time on student behavior. This has not yet been seen in the published literature; creating a standard may create awareness of this possibility and open developers towards having AIS systems work together to support students in engaging more effectively with learning content.

CONCLUSIONS AND RECOMMENDATIONS

We propose to create a standard for representing disengaged and engaged behavior, at a high level, within adaptive instructional systems. This standard will represent behavior as a set of categories, with sub-categories as appropriate; AIS systems that detect these behaviors will represent the behaviors in terms of these standard categories. Doing so will enable AIS systems to communicate with other AIS systems in a consistent fashion about whether a student is demonstrating a behavior, even if the two systems themselves detect the behavior in very different fashions (or even if one of the systems is itself unable to detect the behavior).

REFERENCES

Aleven, V., Mclaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence in Education*, *16*(2), 101-128.

Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Meheranian, H., Fisher, D., ... & Woolf, B. P. (2007). Repairing disengagement with non-invasive interventions. In *Proceedings of the International Conference on Artificial Intelligence and Education,* 195-202.

Baker, R.S., Corbett, A.T., Koedinger, K.R. (2004) Detecting Student Misuse of Intelligent Tutoring Systems.*Proceedings of the 7th International Conference on Intelligent Tutoring Systems*, 531-540.

Baker, R.S., Corbett, A. T., Koedinger, K. R., Evenson, S. E., Roll, I., Wagner, A. Z., Naim, M., Raspat, J., Baker, D. J. & Beck, J. (2006). Adapting to when students game an intelligent tutoring system. In Proceedings of the 8th International Conference on Intelligent Tutoring Systems (pp. 392-401).

Baker, R.S., Rossi, L.M. (2013) Assessing the Disengaged Behavior of Learners. In Sottilare, R., Graesser, A., Hu, X., & Holden, H. (Eds.) *Design Recommendations for Intelligent Tutoring Systems -- Volume 1 -- Learner Modeling*. U.S. Army Research Lab, Orlando, FL, pp. 155-166.

Bouchet, F., Harley, J. M., & Azevedo, R. (2016). Can Adaptive Pedagogical Agents' Prompting Strategies Improve Students' Learning and Self-Regulation?. In *International Conference on Intelligent Tutoring Systems* (pp. 368-374).

Cocea, M., Hershkovitz, A. & Baker, R. S. J. d. (2009). The impact of off-task and gaming behaviors on learning: Immediate or aggregate? In *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 507-514).

Paquette, L., Baker, R.S., de Carvalho, A., Ocumpaugh, J. (2015) Cross-System Transfer of Machine Learned and Knowledge Engineered Models of Gaming the System. *Proceedings of the 22nd International Conference on User Modeling, Adaptation, and Personalization*, 183-194.

Pardos, Z.A., Baker, R.S., San Pedro, M.O.C.Z., Gowda, S.M., Gowda, S.M. (2014) Affective states and state tests: Investigating how affect and engagement during the school year predict end of year learning outcomes. *Journal of Learning Analytics, 1* (1), 107-128.

Rowe, J., McQuiggan, S., Robison, J. & Lester, J. (2009). Off-task behavior in narrative-centered learning environments. In *Proceedings of the 14th International Conference on Artificial Intelligence and Education* (pp. 99-106).

Sabourin, J., Rowe, J., Mott, B. & Lester, J. (2011). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In *Proceedings of the 15th International Conference on Artificial Intelligence in Education* (pp. 534-536).

San Pedro, M.O.Z., Baker, R.S.J.d., Bowers, A.J., Heffernan, N.T. (2013) Predicting College Enrollment from Student Interaction with an Intelligent Tutoring System in Middle School. *Proceedings of the 6th International Conference on Educational Data Mining*, 177-184.

Shih, B., Koedinger, K. R., & Scheines, R. (2011). A response time model for bottom-out hints as worked examples. *Handbook of educational data mining*, 201-212.

Tscholl, M., Biswas, G., Goldberg, B. S., & Sottilare, R. A. (2016). Detecting Metacognitive Strategies through Performance Analyses in Open-Ended Learning Environment. In *Proceedings of the 24th International Conference on Computers in Education*.

ABOUT THE AUTHORS

Dr. Ryan S. Baker *is Associate Professor at the University of Pennsylvania and Director of the Penn Center for Learning Analytics. He was the founding President of the International Educational Data Mining Society, has taught four MOOC instances, and was the co-developer of the Baker Rodrigo Ocumpaugh Monitoring Protocol. Baker has co-authored published papers with over 250 colleagues.*

Chad Coleman *is a doctoral student at Teachers College Columbia University and a participating member of the Penn Center for Learning Analytics at the University of Pennsylvania. His research interests center around the intersection of human computer interaction, instructional systems, and learning analytics.*