

Adding eye-tracking AOI data to models of representation skills does not improve prediction accuracy

Martina A. Rau
Department of Educational Psychology
University of Wisconsin—Madison
1025 W. Johnson St
Madison, WI 53706
+1-608-262-0833
marau@wisc.edu

Zach Pardos
2nd author's affiliation
1st line of address
2nd line of address
Telephone number, incl. country code
2nd E-mail

ABSTRACT

Visual representations are ubiquitous in STEM instruction. Representation skills allow students to use visual representations to learn about concepts. It seems reasonable to hypothesize that we can gather useful information about representation skills from eye-tracking AOI data that assesses how students pay attention to representations. We tested this hypothesis by comparing cognitive models with and without eye-tracking AOI data. Specifically, we used Bayesian Knowledge Tracing and Long Short Term Memory models. We evaluated these models based on their accuracy in predicting students learning of knowledge components that assess representation skills. Eye-tracking AOI data did not improve the prediction accuracy of our cognitive models. We compare our results to prior research to generate hypotheses for future research.

Keywords

Visual representations, intelligent tutoring system, eye-tracking, Bayesian Knowledge Tracing, Long Short Term Memory models.

1. INTRODUCTION

STEM instruction typically uses visual representations that depict to-be-learned content [1]. To learn content knowledge, students have acquire *representation skills*: the ability to use visual representations to learn [2]. Instructional support is most effective if it not only focuses on students' learning of content knowledge, but also on their learning of representation skills [1]. Intelligent tutoring systems (ITSs) have the capability to adapt to the individual student's needs [3]. They do so based on a cognitive model that infers the student's knowledge level based on interactions with the ITS [3]. Hence, the goal of cognitive modeling is to accurately model students' learning in real time [4]. A limitation of this research is that it has mostly focused on students' content knowledge, not on representation skills.

It seems reasonable to assume that we can gather useful information about students' learning of representation skills from their visual attention to representations [5]. However, most prior eye-tracking research involved relatively simple learning materials; typically expository text paired with one additional visual representation. By contrast, ITSs are more complex. Second, prior research has not focused on using eye-tracking AOI data to model students' learning of representation skills. For example, Conati's research group used eye-tracking data in cognitive models found that it can improve predictions of students' learning of content knowledge [6]. This paper tests the hypothesis that eye-tracking AOI data improves cognitive models.

2. DATASET

We used data from a lab experiment that collected students' eye-tracking data while they worked with an ITS for chemistry for 3h [7]. 117 undergraduates participated in the experiment. For our

analyses, we used log data from the ITS and eye-tracking data. To analyze the log data, we constructed a knowledge component (KC) model that relates each problem-solving step to the underlying skill. KCs corresponded to representation skills. To analyze the eye-tracking data, we generated visual attention features that assess how students process the visual representations with areas of interest (AOIs) that correspond to the representations. We also created AOIs for the parts of the screen where students solve problems, for the hint window, and for the periodic table that students could show and hide. We included only logged events and first attempts that were tagged with a KC with more than 30 data points. Our final dataset comprised a total of 30,893AOI and log events.

3. ANALYSES

We used two cognitive modeling approaches: Bayesian Knowledge Tracing (BKT) and Long Short Term Memory (LSTM) models. Both analyses used a 5 fold cross validation scheme which was created by assigning students to folds once.

BKT is the standard cognitive modeling procedure in research on ITSs [8]. We used BKT to evaluate a cognitive model representing performance prediction based on a student's history of incorrect and correct responses to questions of the same knowledge component. Following standard practice, we evaluated different guess and slip equivalence classes, which included using a different guess and slip per problem or per step. In previous work [9], separate guess and slip classes at the problem level resulted in a 10% gain in accuracy on ITS dataset. We applied this model to KCs without eye-tracking AOI data and to a version with eye-tracking AOI data. For the latter model, we fit a separate learning rate for each AOI within a problem.

All BKT models were fit with expectation maximization (EM) with max iteration of 100 and epsilon of 1e-6 as stop criteria. The best models in terms of log-likelihood used 40 EM restarts with initial parameter values. For prior these were drawn from a uniform random distribution, while the values for learn, guess, and slip were capped at 0.40, 0.40, and 0.30 respectively.

LSTM models are a subset of Recurrent Neural Networks (RNN). Recent progress in image classification with convolutional neural networks utilizes its ability to learn features that have more predictive power than manually crafted features (e.g., edge detection), previously the state of the art for image classification. In a similar vein, we used LSTM so that features of eye-tracking AOI data not yet known to be important could potentially be picked up. Therefore, the LSTM in represents a powerful detector to find out if there is a useful predictive signal in our sequences of eye-tracking AOI data.

We used two LSTM variants on RNNs that add a state to the hidden layer called the cell state which allows the network to

more effectively remember actions that occurred in the past when piecing together patterns in sequential input. We compared versions that utilized eye-tracking AOI data to versions that did not. Both LSTM models utilized the identical amount of information as their BKT with-eye and without-eye data counter parts and both trained a separate model per KC. In the case of LSTM models; eyeHeader, problemID-AOI, and Outcome comprised the feature vector. In both LSTM models, there is an instance of training data for every response given by a student. While non eye-tracking models were trained on sequence lengths that extend as long as the longest response sequence, AOI sequences were limited to the most recent N events, where N was defined as the maximum number of responses of any student in the training data + the median number of AOI events per student. This was done so that the data could fit into memory using 8bit signed integer matrices on a single large memory compute node.

4. RESULTS

After the 5 fold cross validation, RMSE was calculated per student. For a baseline reference, the RMSE of predicting the average percent correct for each KC was 0.39062. Models without eye-tracking data performed better than all of the models with eye-tracking data. Among the BKT models, problem was the better choice for assigning guess and slips over stepname, agreeing with prior work on ITS data [9]. Among LSTM models, extending the number of training epochs from 5 to 10 resulted in the most substantial gain of any model when not using eye-tracking but more epochs lead to overfit with the eye-tracking model. LSTMs, given the same problem-id and response data, were better able to leverage the information towards prediction accuracy than BKT, although both relied on a KC model. Differences between predictions were statistically reliable ($ps < 0.05$), as determined by a paired t-test of squared residuals between all adjacent models in the list with the exception of the LSTM model with 5 epochs and the BKT model with problem-id as guess/slip, which both used eye-tracking AOI data.

5. DISCUSSION

Our results stand in contrast to our hypothesis: using two cognitive modeling approaches, we did not find evidence that eye-tracking AOI data improves the accuracy of the model's prediction. This finding is noteworthy for the following reasons. First, it is counterintuitive because we tend to assume that visual attention is an important factor in assessing representation skills. Second, our finding stands in contrast to prior research on learning with text paired with one additional visual representation, where students view rather than interact with the material. The difference between prior work and our work is that our study used a complex learning environment, where students manipulated visual representations to solve problems. Third, our results stand in contrast to prior work, which found that eye-tracking AOI data can improve the accuracy of cognitive models of students' learning of content knowledge. The difference between prior work and our work is that our cognitive model assessed students' learning of representation skills, which reflects students' knowledge about the content and about visual representations.

One possible explanation is that prior eye-tracking research on learning with simple materials did not assess whether eye-tracking AOI data adds predictive accuracy to log data—because these materials do not generate log data. Second, representation skills may reflect not how students inspect visual representations, but how they use information from the representations to solve problems, which is sufficiently captured by the log data—

particularly if the representations themselves are interactive and hence generate log data that can be used in cognitive models. Third, the fact that we modeled representation skills rather than content knowledge may explain why our results stand in contrast to prior work by Conati's group. We used a KC model that was specifically designed to assess students' representation skills. Even if eye-tracking AOI data assesses representation skills, it may simply not improve the accuracy of our cognitive model because the KC model already captures this information.

A limitation of our research results from the fact that the granularity of our AOIs was fairly coarse. Subtle cognitive signals may exist at fine grained resolutions which may require diving into the raw eye-tracking AOI coordinates. A second limitation was the exploration of hyper parameters. While this is always a caveat of any analysis using machine learning, a particular set of hyper parameters may exist which unlocks the predictive utility of the existing eye-tracking AOI data.

In sum, our findings suggest that eye-tracking AOI data does not necessarily add information relevant to students' representation skills, compared to what can be captured by a well-crafted KC model of representation skills. This rationale amounts to a new hypothesis that should be tested in future research: namely that adding representation skills to cognitive models of content knowledge may improve prediction accuracy in the same way as the addition of eye-tracking AOI data would.

6. ACKNOWLEDGMENTS

This work was supported by a grant from the National Science Foundation (IIS: BIGDATA 1547055).

7. REFERENCES

- [1] Gilbert, J.K.: 'Visualization: An emergent field of practice and inquiry in science education': 'Visualization: Theory and practice in science education' (Springer, 2008), pp. 3-24
- [2] NRC: 'Learning to Think Spatially' (National Academies Press, 2006)
- [3] Koedinger, K.R., Corbett, A.: 'Cognitive Tutors: Technology bringing Learning Sciences to the classroom': 'The Cambridge Handbook of the Learning Sciences' (Cambridge University Press, 2006), pp. 61-77
- [4] Baker, R., Siemens, G.: 'Educational Data Mining and Learning Analytics': 'The Cambridge Handbook of the Learning Sciences' (Cambridge University Press, 2014), pp. 253-272
- [5] Mason, L., Pluchino, P., Tornatora, M, Ariasi, N.: 'An eye-tracking study of learning from science text with concrete and abstract illustrations', The Journal of Experimental Education, 2013, 81, (3), pp. 356-384
- [6] Bondareva, D., Conati, C., Feyzi-Behnagh, R., Harley, J.M., Azevedo, R., Bouchet, F.: 'Inferring learning from gaze data during interaction with an environment to support self-regulated learning': 'Artificial Intelligence in Education' (Springer, 2013), pp. 229-238
- [7] Rau, M.A., Wu, S.: 'ITS support for conceptual and perceptual processes in learning with multiple graphical representations': 'Artificial Intelligence in Education' (Springer International Publishing, 2015), pp. 398-407
- [8] Anderson, J.R., Boyle, C.F., Corbett, A.T., Lewis, M.W.: 'Cognitive modeling and intelligent tutoring' (Elsevier Science The MIT Press, 1990)
- [9] Pardos, Z.A., Heffernan, N.T.: 'KT-IDEM: Introducing Item Difficulty to the Knowledge Tracing Model': 'User Modeling, Adaption and Personalization' (Springer, 2011), pp. 243-254