Towards Better Understanding of Transfer in Cognitive Models of Practice

MICHAEL V. YUDELSON, PHILIP I. PAVLIK JR., and KENNETH R. KOEDINGER, Carnegie Mellon University

Achieving transfer – the ability to apply acquired skills in contexts different from those contexts the skills were mastered in – is, arguably, the sine qua non of education. Capturing transfer of knowledge has been addressed by several user modeling and educational data mining approaches (e.g., AFM, PFA, CFA). While similar, these approaches use different underlying structures to model transfer: Q-matrices and T-matrices. In this work, we compare of a more traditional Q-matrix-based method and the relatively new and more complex T-matrix based method. Comparisons suggest that the T-matrix, although demonstrating only marginally better fits, offers a more interpretable and consistent picture of learning transfer.

Transfer is, arguably, the sine qua non of education, in which the primary goal for the learner is to be able to apply new skills in contexts often different from the ones they were mastered in. Because of this, achieving transfer from the math skills trained by computer-aided educational systems to the math abilities used later in a student's life is a litmus test for determining the success of such systems as a whole.

Educational systems access the transfer by tracking students' learning with help of some model (usually math based). This model is also used to make instructional decisions. It model can be simple or complex, but if it does not provide practice with transfer in mind, it is unlikely that long term-learning will be strong. In psychology, for example, it is well known that what optimizes immediate performance is unlikely to be the practice that optimizes long-term retention or transfer [Schmidt and Bjork 1992]. Due to this paradox of "desirable difficulties" an educational software system needs a model that is clever enough to see not only the effect of practice on repetition, but also see the effect of practice on transfer. This sort of model is unlikely to be easy to configure, since understanding transfer is paramount to understanding education itself, because of how crucial transfer is to a flexible education.

This problem has been addressed by the user modeling and educational data mining communities. While many approaches have been attempted over the years to handle transfer, the Q-matrix method of assigning latent knowledge components (KCs) to particular problems or problem steps has begun to have a large following , even being supported with a suite of logging and analysis tools [Koedinger et al. 2008]. Q-matrix (or question matrix) methods have a long history [Birenbaum et al. 1992; Tatsuoka 1983]. They assign rules to questions so as to determine the aggregate difficulty of items in an instructional situation. The Q-matrix method is notable for the way it specifies these rules abstractly as latent variables that are contained in one or more questions. In contrast, a newer "T-matrix" method assumes learning is less abstract by not describing latent variables, but rather looking at the transfer effects with a question by question matrix where each question causes learning that effects other questions directly [Pavlik Jr. et al. 2011].

In order to better understand transfer, we decided to meticulously compare these two approaches. This contrast is interesting, because the Q-and T-matrix models make clearly different assumptions about transfer.

This work is supported by the U.S. Department of Education (IES-NCSER) #R305B070487 and was also made possible with the assistance and funding of Carnegie Learning Inc., the Pittsburgh Science of Learning Center, DataShop team (NSF-SBE) #0354420 and Ronald Zdrojkowski.
Author's addresses: M. V. Yudelson, P.I. Pavlik, and K.R. Joedinger, Human Computer Interaction Institute, School of computer Sciences, Carnegie Mellon University.

In the Q-matrix case, the method specifies transfer by the sharing of latent construct between 2 items. Because of this sharing of the latents, it is seldom the case that transfer is asymmetric in the model. On the other hand, in the T-matrix case, there is no sharing of latents. Instead, the T-matrix - a more complex method - specifies each directional pairwise relationship between a transferred-from item and a transferred-to item individually. Therefore, the T-matrix model is more fit for capturing asymmetry of transfer.

If we trace these assumptions of Q-and T-matrix-based models to the learning science and psychology literature, we can see a problem that inspires our comparison. Specifically, there are notable examples in the literature where asymmetrical transfer occurs strongly (e.g.[Bassok and Holyoak 1989]). Similarly, we find cases where learning is optimal for one condition (e.g. concrete task), but when transfer is analyzed, another condition is more beneficial (e.g. a more abstract task) [Goldstone and Son 2005]). However, there exist other examples of transfer being successfully modeled as latent skills (rf. [Singley and Anderson 1989]). Therefore, the question is, given the prior success of the Q-matrix-based methods and the potential benefits of a relatively new T-matrix method, can we find the evidence that warrants coping with T-matrix complexity for the benefit of its greater flexibility?

In this work we attempt to answer this question by comparing Q-matrix models and T-matrix models to determine which fit the data better and which provide a richer qualitative model of the data. This comparison allows us to establish whether the added complexity of the T-matrix method is justified. While the overall fit is important, we also keep in mind that the goal of educational data mining is not just to produce the optimal model, but also to produce a model that has educational implications. Because of this, we consider the different practical implications of each model to help us establish which one is qualitatively more useful for understanding the educational data.

Our comparison of user modeling methods based on Q-matrix and T-matrix suggest that the latter, although demonstrating only marginally better fits, offers a more interpretable and consistent picture of learning transfer. The results support the claim that in specific cases there is a need for the types of analysis where asymmetric transfer is specifically modeled. Despite the fact that T-matrix-based models are not quantitatively better overall, we found that there were consistent qualitative patterns of asymmetric transfer revealed by the T-matrix model. By comparing the model parameters that governed transfer in each model we were able to see that often Q-matrix-based method were producing bidirectional transfer models by averaging. At the same time the T-matrix-based model was able to show asymmetry clearly.

REFERENCES

Bassok, M. and Holyoak, K. 1989. Interdomain transfer between isomorphictopics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition 15*, 153–166.

Birenbaum, M., Kelly, A., and Tatsuoka, K. K. 1992. *Diagnosing knowledge states in algebra using the rule space model Educational Testing Service.* NCTM, Princeton, NJ.

Goldstone, R. and Son, J. 2005. The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences 14*, 69–110.

Koedinger, K., Cunningham, K., Skogsholm, A., and Leber, B. 2008. An open repository and analysis tools for fine-grained, longitudinal learner data. In *Proceedings of the 1st International Conference on Educational Data Mining*, R. Baker and J. Beck, Eds. 157–166.

Pavlik Jr., P. I., Yudelson, M. V., and Koedinger, K. R. 2011. Using contextual factors analysis to explain transfer of least common multiple skills. In *Proceedings of the 15th International Conference on Artificial Intelligence in Education (AIED 2011)*, J. Kay, S. Bull, and G. Biswas, Eds. Springer-Verlag, New York, NY (accepted).

Schmidt, R. and Bjork, R. 1992. New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science 3*, 207–217.

Singley, M. and Anderson, J. 1989. *Transfer in the ACT* theory. In The transfer of cognitive skill.* Harvard University Press VIII, Cambridge, MA.

Tatsuoka, K. 1983. Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement 20,* 4, 345–354.