CHAPTER 3 – Important Considerations for Learner Models: Transfer Potential and Pedagogical Content Knowledge

Alan Lesgold[1] and Arthur Graesser[2]
[1] University of Pittsburgh; [2] University of Memphis

Introduction

In the age of the intelligent machine, training of humans inevitably requires training for transfer. If we knew exactly what a person should do in a situation, we could program a computer or robot to do it instead. However, no two situations are exactly the same so uncertainty and differences arise in matches between the current situation and episodes of the past. Modeling transfer has traditionally meant getting as close as possible to modeling the total expertise trainees were expected to have in handling episodes of the past and future. We suggest that there may be more effective approaches today. One possibility is to model knowledge (general or specific) that would permit a person to represent a situation in a form that can be handled by diverse training and transfer trajectories. A second possibility would be to establish a principled way to prioritize which knowledge should be overlearned so that it can be "stretched" to novel situations not covered in training. A related approach would be to do both. For example, one could perform situation modeling within the domain of interest along with cognitive load analyses of the situation modeling examples drawn from the target domain. This chapter explores these options and considers what task analysis and expert modeling is needed to capture them.

If we want people to carry out performance of a task in a precise way, we can conduct a task analysis and then teach each of the elements of the performance and directly measure how well each element is acquired. This approach to training worked remarkably well for the training of line workers in the industrial age and of enlisted personnel in the military prior to the knowledge revolution (Collins & Halverson, 2009). It assumed availability of supervisors or officers who would creatively handle unexpected challenges and provide direction that allowed teams to adapt to emergent situations. It was generally assumed that these leaders did not need carefully monitored complete training for their roles but rather that they would be selected as being intelligent enough to prove a useful bridge from the formal training of their subordinates to the actual situations their unit might encounter.

Now that we are entrenched in the knowledge revolution, however, routine performance can be specified so precisely that we can teach a person or program an intelligent machine to perform these procedures competently. Moreover, once empowered by intelligent tools, every worker or soldier has a role akin to the leaders or officers of times past. Nevertheless, experience has shown that most workers in intelligent work environments need substantial training, even if we cannot fully drill them on every aspect of the performances we hope they will exhibit. This creates a need to consider transfer in the design of training. That is, we are no longer merely preparing people by teaching them all the elements, rules, or procedures that define perfect performance in every likely circumstance. Rather, they need to be prepared for emergent situations that deviate from original training. Our hope is that our trainees will show transfer from the specific prescribed training to ideal performance on emergent tasks.

Related Research

The profound challenge is that there typically is unspectacular transfer from training to new situations (Banich & Caccamise, 2010). A classical study by Hayes and Simon (1977) had college students attempting to solve a series of problems that had structurally identical solutions but varied in surface characteristics, such as substituting names of characters and objects (Hobbits and Orcs vs. Monsters and

Globes). Transfer between four successive problems was near zero unless students were explicitly instructed to make similarity connections between the problems. Gick and Holyoak (1980) investigated whether the structure of a story about troops converging on a location would help college students solve a radiation problem that had a direct structural isomorph to the story. The transfer was modest without cognitive activities that intentionally tried to establish correspondences between the two representations. At a more global practical level, the order of courses in a curriculum is rarely backed by substantiating transfer data (Vuong, Nixon & Towle, 2011). For example, there is not an abundant body of empirical evidence that calculus helps engineering or that physics helps students understand chemistry. Knowledge and skills are highly constrained by specific characteristics of the subject matter.

The above examples present a dismal picture of transfer, but it can be countered by exemplars of successful transfer. The literature is replete with evidence that training on particular tasks can facilitate performance on similar tasks in the future. The devil is in the details of the similarity of the stimuli, tasks, and associated cognitive representations (Gentner & Markman, 1997).

The transfer problem has enormous implications for the GIFT architecture (Sottilare, Brawner, Goldberg & Holden, 2012), particularly with respect to the distinction between domain-independent and domain-dependent modules. The current GIFT architecture specifies that the sensor and pedagogical modules are domain-independent whereas the subject-matter knowledge is domain dependent. The implication of the domain-independent modules needs to be clarified. Does that mean that there should be significant transfer between tasks that have similar subject matters but different sensor and pedagogical modules? Does that mean that the same pedagogical methods can be applied to a broad set of subject matters, as opposed to the pedagogical modules being distinctively tailored for particular subject matters? Does subject matter (domain knowledge) reign supreme over sensor and pedagogical modules? What are the priorities among these GIFT modules on the matter of transfer?

We propose that a model of student[2] learning has two categories of knowledge in an intelligent training system. First, it knows how well the student has mastered the elements specifically being trained. Second, it knows how prepared the student is to confront categories of situations that cannot be predicted and cannot be rehearsed adequately. This latter requirement means that such systems must embody a theory of transfer that can allow us to know how far a student's knowledge might stretch.

Consider, for example, a football coach. The coach might train the team to execute a particular kind of play, such as an option running play. The team can practice each of the two or three ball carrier options, so that every team member knows who to pass the ball to, who to get the ball from, or where the players position themselves to block the defense. However, the practice sessions only rehearse some of the possibilities because the defensive players are intelligent entities and do not necessarily behave exactly as the sham defense set up during the practice sessions. Yet the coach is hopeful that the team will perform well in every case.

Some of the performance environments are predictable. By having the defense behave intelligently and conform to known football best practices, the coach can create a variety of practice situations that anticipate what will happen in an upcoming game. However, the game will have situations that deviate from the ones on which practice occurred. A theory of practice is useful to the extent that it allows the coach to predict game performance from practice performance, select the right practice situations, and decide when there has been enough practice.

---

[2] For simplicity of exposition, we use terms like "student model" and "students" even though the primary audience for this volume are training developers and training system designers, not schoolteachers or professors.

Unpredictable situations will nevertheless occur. It might be icy or rainy. The field might be muddy or not the same surface as that on which practice occurred. The players may change as wounds heal, injuries occur, and so on. Consequently, there need to be global indicators of whether team members are ready for a wider range of unpredictable occurrences. A decision is needed as to whether the current situation is under the realm of status quo or an unusual case. An adequate model would identify specific performances that a player can do well in prototypical situations as well as some estimates of which categories of unpredictable cases can be expected to be adequately handled. The latter models the transfer potential of what was explicitly learned. One way to address the problem of transfer is to consider the two skill sets separately. That is, there are the specific performance skills that encompass expertise in the target domain. There are also *situation representation* capabilities that allow a student to classify a novel situation as one in which particular learned performances can be successful. These situation representational skills mediate the availability of learned rules to be applied in emergent situations. Our experiences in building intelligent training systems have convinced us that these two skill sets should be handled in different ways.

The first author has developed and tested intelligent training systems for maintaining equipment in the military and business sectors, including troubleshooting equipment failures (Gott & Lesgold, 2000; Gott, Lesgold & Kane, 1997; Lesgold & Nahemow, 2001). For the specific performance skills, there was the standard approach of building an expert model that was capable of solving the range of tasks that was targeted for training. The initial assumption was that once each rule in the expert model was demonstrated by the student, training was considered completed. However, our experience quickly revealed that this was not entirely the case because the target domains we addressed were extremely complex. All of the information needed to trigger each of the rules relevant for a given task was present in the task domain, but there was so much information that a serious challenge remained in representing the situation at hand sufficiently to trigger appropriate rules.

The challenge of system complexity required us to train the situation representation skills needed to perceive the problem domain adequately to trigger expert rules. Our expert informants initially believed that this required training on recognizing the various components of the complex target domain. For example, in the case of avionics test stations, this would include the system modules and the components of the modules. However, it was apparent that all of the students could recognize all of the modules and their components, even at the beginning of training. The needed situation representation skills were more subtle than our expert informants realized.

A reasonable training method would be to directly teach how to recognize which aspects of a situation are important. This is done in football when players are taught the names and overall strategies behind various types of plays. It also is done in medicine when physicians in training directly learn to recognize various syndromes for which diagnostic rule sets are known. This approach to teaching for transfer can work when there are recurrent examples of these various plays or syndromes. In football, that happens because the coach has studied game films and knows the patterns that the opposing team is likely to use. In medicine, it happens because various genetic and environmental factors predispose human bodies to exhibit various syndromes or patterns of malfunction (e.g., lots of us overeat and under-exercise, with the bodily response being pretty stereotypic).

Engineered systems generally are built and modified to adapt to the range of circumstances in which they are deployed, but unfortunately they often do not have predictable patterns of breakdown. If those existed, they would be engineered away. This makes it harder to teach situation representation by teaching how to recognize syndromes; there are few if any recurrent syndromes to teach. This required us to take a different approach that did not explicitly teach students to watch for specific patterns, at least for the most part. Instead, students were provided a useful *range* of experiences that prompted them to construct rulebased knowledge on top of their prior experiences with the task domain. This more flexible emphasis on the range of the situations was expected to extend the students' knowledge to new situations.

The scheme that was developed was called *intelligent coached apprenticeship*. Basically, we pushed students to go beyond their overlearned knowledge to solve novel problems (Gott & Lesgold, 2000). We expected that none of the problem tasks our training system provided would be solvable by the students without help. Five levels of help were available. The lowest level listed the sequence of actions the student already had taken to address the problem. More than half of the time, that level of hint was sufficient to keep the student moving toward a solution to the problem. If the student was still stuck, there would be more explicit levels of hint (e.g., where to look for information) until a final hint level that told the student exactly what to do on the next step. For the population our training addressed, the desire to learn was strong, and students seldom asked for more help than they needed. After completing the solution for the problem, they also were able to study a comparison of their solution path to that of an expert.

The central hypothesis in this research (Gott & Lesgold, 2000; Gott, Lesgold & Kane, 1997) is that scaffolding for students to shape their own construction of situation representation skills would produce the needed development of those skills, even if we could not explicitly list them all and even if it was not feasible to build a complete set of situation representation rules and train on each. It was reassuring that our efforts were highly successful. We had expert technicians develop a collection of far transfer tasks that involved applying the expert rules we taught to new hardware. The transfer hardware was a fictitious piece of hardware that could still be diagnosed using the expert rule set that our training system embodied. The students we trained performed in about the same range as experts on those problems (Gott & Lesgold, 2000; Gott, Lesgold & Kane, 1997).

It may eventually be possible to teach the needed situation representation skills directly rather than to assume them and indirectly stimulate their construction by the students. For example, Forbus et al. (2007) have made progress in specifying what a computer needs in order to learn similar kinds of representational skills. Nevertheless, the scheme we developed is worthy of further exploitation. On the one hand, it focuses on the expert rules needed to do the necessary problem solving, and on the other hand, it does not ignore the reality that some level of further knowledge construction is needed if those rules are to be available when needed in addressing novel tasks.

There will always be uncertainty that an individual student's understanding of the task domain embodies entirely the same constructs that the training designer might have had. This is especially the case in complex technical domains, where the training designers often know much more basic science than the technicians being trained. As an example, consider the terms used in expert rules for diagnosing failures of an ion beam system for writing circuits on computer chips (Lesgold & Nahemow, 2001). A complete account requires knowledge of quantum physics, silicon chemistry, and optics. Technicians two years out of high school could learn the diagnosis rules well enough to apply them in transfer situations using the training technique of providing difficult problems with scaffolding and post-problem reflection (Lesgold & Nahemow, 2001).

Discussion

The intelligent coached apprenticeship was impressively effective, producing learning effect sizes exceeding one standard deviation. These results are on par with or exceed intelligent tutoring systems that have been developed and tested during the last decade (Graesser, Conley & Olney, 2012; VanLehn, 2011). However, questions remain about the information that is needed in the students' situation representation to assure adequate learning. There are also questions directly relevant to GIFT (Sottilare et al., 2012).

The systems described in the previous section did not log information about the students' situation representation, and it did not retain its short-term estimations of student reasoning strategies. Instead, the systems were validated by conducting far transfer studies. This is an expensive way to proceed. It is worth considering simpler ways to assure adequate learning of rules that provide substantial transfer. Some of these possibilities are presented in this section.

One important step would be to record information about the level of coaching/scaffolding that a student requires. The systems described above did that, but the data were not retained beyond their presence in a post-problem recapitulation of the student's activity and a comparison of that to an expert solution. GIFT, in contrast, routinely records in the log files the raw sensor data, computer-student interactions, pedagogical strategies being implemented, expected student responses (both correct and misconceptions), actual student responses, and other data that can be mined for system improvement. At any given moment in solving a problem in the intelligent coached apprenticeship systems, the system kept track of what the best next step for a student. Therefore, it would be straightforward to record data that indicates that the student has a misconception or does not know what to do next (see, for example, the inference of "malrules" by systems VanLehn [1981] has built). Informative episodes occur when the student asks for coaching or performs a non-optimal next step. As the student starts taking appropriate actions, the past records of incomplete learning could be retired and eventually there would be confidence that a rule has been learned broadly. GIFT supports such data mining and machine learning activities.

Note that the approach is substantially different from direct training of a rule. In direct training, the student is placed in circumstances where the rule should be triggered and then taught explicitly to apply it. The problem with the direct training approach is that the rules flagged as learned may not be triggered in circumstances where complex tasks are being performed without attention focused directly on possible circumstances where the rule applies. We have known for almost a century (Whitehead, 1929) that specifically learned bits of knowledge often are not used in broader circumstances where they should apply. By focusing more of learning on explicitly stretching one's knowledge, this problem has been avoided. The question arises whether training of humans is needed at all, given that the intelligent coached apprenticeship systems relied upon expert models to provide coaching. Expert models would perhaps not be able to transfer to novel situations any more than novice technicians. We were able to formally represent each problem in a way that permitted the expert system to solve it, but those representations were only possible because of the additional knowledge of the experts who developed them. That more implicit knowledge is exactly what intelligent coached apprenticeship endeavors to train.

Decisions will need to be made on how the above mechanisms in the intelligent coached apprenticeship would be implemented in GIFT. The student-constructed situation representation is not merely stored in the GIFT Domain Module, but rather is apparently derived in the Learner Module from a combination of activities involving the Sensor Module, Pedagogical Module, the Domain Module, and history of the logged data. Does this complex interactivity clash with the modular assumption that differentiates the domain-dependent Domain Module from the remaining domain-independent components?

Another important step for implementing effective intelligent coached apprenticeship systems resides in tracking mastery of rules in broad contexts. The problem sets presented to students need to be sufficiently challenging and span a wide enough range of situations. They can never span all the situations a student will encounter when applying what is being taught. However, they need to span a sufficiently wide range that they force students to reflect on why each rule is applicable and the range of possible situations of applicability. For the two rather different domains developed by Lesgold and his colleagues, the expert technicians helped build problems that are sufficiently diverse that there was transfer to novel situations.

One situation involved diagnosis of failures in complex switching systems designed to test electronic components from military aircraft. The range of problems experts helped us develop was sufficient to produce far transfer performance that was about as good as that of domain experts (Gott, Kane & Lesgold, 1995; Gott, Lesgold & Kane, 1997). The second domain in which we did similar training and testing with similar results was the diagnosis of failures in machines used to make computer chips, notably ion deposition systems that put layers on chips and ion beam implant systems that write circuits on those layers (Lesgold & Nahemow, 2001). The switching systems had underlying knowledge associated with simple electrical properties of circuit continuity and magnetics (relays were involved). The chip-making systems involved underlying knowledge of electrical systems, heat distribution, silicon gas compounds, movements by robots, and to some extent, even more complex basic science.

The same basic approach worked for both domains so it is at least a reasonable conjecture that it can work much more widely. It would be worthwhile, however, to study the utility of logging indications of uncertainty, lack of knowledge, or misconceptions at points where the rules in an expert system are not applied routinely or accurately. Such information might allow better decisions on how much training is needed to produce mastery. Once again, the logged data, data mining, and machine learning facilities of GIFT are, in principle, equipped to implement this approach. Decisions will need to be made on the division of labor among GIFT modules when accommodating changes to the system.

Recommendations and Future Research

The intelligent coached apprenticeship system described in this chapter underscores the importance of capturing situation representations that are constructed by the students in order to handle a diverse range of cases applicable to transfer situations. Domain experts will be needed to select the problems that deviate from routine cases that can be trained explicitly. Research will therefore be needed to understand how these cases/problems are selected, the mapping between training and transfer cases, the situation representations that students construct, and their ability to identify unusual cases. A detailed analysis of the log data should be helpful in these research efforts and also in modification of the systems in iterative development.

There are two major recommendations directly relevant to GIFT. First, decisions will need to be made on how the specific modules will participate in the intelligent coached apprenticeship system. It is not a simple matter of storing content in the Domain Module. Second, decisions will need to be made on how to represent the information stored in the log files and the various modules. It is not a simple matter of storing everything. The features, content, and structures will need to be able to support new cases and situation representations in addition to domain-dependent and domain-independent information.

References

Banich, M.T. & Caccamise, D. (2010)(Eds.) Generalization of knowledge: Multidisciplinary perspectives. New York: Psychology Press. .

Collins, A. & Halverson, R. (2009). Rethinking education in the age of technology: The digital revolution and schooling in America. New York: Teacher College Press.

Forbus, K., Riesbeck, C., Birnbaum, L., Livingston, K., Sharma, A. & Ureel, L. (2007). Integrating natural language, knowledge representation and reasoning, and analogical processing to learn by reading. Proceedings of AAAI-07: Twenty-second Conference on Artificial Intelligence. (1542-1547). Vancouver, BC: AAAI.

Gentner, D. & Markman, A. (1997). Structure mapping in analogy and similarity. American Psychologist, 52, 45-56.

Gick, M.L. & Holyoak, K.J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Gott, S. P., Kane, R. S. & Lesgold, A. (1995, February). Tutoring for transfer of technical competence. Armstrong Laboratory Technical Report AL/HR-TP-1995-0002. Brooks AFB, TX: Air Force Materiel Command.

Gott, S. P., Lesgold, A. M. & Kane, R. S. (1997). Promoting the transfer of technical competency. In S. Dijkstra, et al. (Eds.), Instructional design: International perspectives (vol. 2). (221-250). Hillsdale, NJ: Erlbaum.

Gott, S. P. & Lesgold, A. M. (2000). Competence in the workplace: How cognitive performance models and situated instruction can accelerate skill acquisition. In R. Glaser (Ed.), Advances in instructional psychology. (239-327). Hillsdale, NJ: Erlbaum.

Graesser, A.C., Conley, M. & Olney, A. (2012). Intelligent tutoring systems. In K.R. Harris, S. Graham, and T. Urdan (Eds.), APA Educational Psychology Handbook: Vol. 3. Applications to Learning and Teaching (pp. 451-473). Washington, DC: American Psychological Association.

Hayes, J.R. & Simon, H.A. (1977). Psychological differences among problem isomorphs. In J. Castellan, D.B. Pisoni & G. Potts (Eds.), Cognitive theory, vol. 2. Hillsdale, NJ: Erlbaum.

Lesgold, A. & Nahemow, M. (2001). Tools to assist learning by doing: Achieving and assessing efficient technology for learning. In D. Klahr & S. Carver (Eds.), Cognition and instruction: Twenty-five years of progress. Mahwah, NJ: Erlbaum.

Sottilare, R.A., Brawner, K.W., Goldberg, B.S. & Holden, H.K. (2012). *The Generalized Intelligent Framework for Tutoring (GIFT)*. Orlando, FL: U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED).

VanLehn, K. (1981). Bugs are not enough (Tech. Rep. No. CIS-11). Palo Alto, CA: Xerox Research Center.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems. Educational Psychologist, 46, 4, 197-221.

Vuong, A., Nixon, T. & Towle, B. (2011). A method for finding prerequisites within a curriculum. In Pechenizkiy, M., Calders, T., Conati, C., Ventura, S., Romero , C., and Stamper, J. (Eds.), Proceedings of the 4th International Conference on Educational Data Mining (pp. 211-216).

Whitehead, A. N. (1929/1967). The aims of education and other essays. New York: The Free Press.