Transferring Primitive Elements of Skill Within and Between Tasks

Logan J. Gittelson (gittel@rpi.edu)
Cognitive Science Department, 110 8th Street
Troy, NY 12180 USA

Niels A. Taatgen (n.a.taatgen@rug.nl)
Department of Artificial Intelligence, Nijenborgh 9
Groningen 9747 AG Netherlands

Abstract

The primitive elements of skill theory proposes a set of approximately 2000 *primitive information processing elements* (PRIMs) (Taatgen, 2013) that compose all cognitive acts by combining and recombining to produce learning and transfer. By this theory, learning is transfer and transfer results from learning as the primitive elements combine to form new elements based on task demands and these more complex elements are reused later in learning (thereby producing increase in skill) and repurposed by different tasks (thereby producing transfer). We illustrate PRIMs in this paper by producing two models of the Balance Beam Task (BBT) and of the Take the Best heuristic (TTB). Although BBT and TTB do not, on the surface, possess much in common, when run in a transfer paradigm (TTB-to-BBT or BBT-to-TTB) each model harvests PRIMs created by its predecessor, thereby demonstrating positive transfer.

Keywords: Balance Beam; Take the Best; ACTransfer; primitive information processing elements (PRIMs)

Introduction

For decades cognitive scientists have been studying the amazing phenomenon of learning. Many researchers have developed theories and models of skill acquisition, but none have fully captured the mechanism behind learning. This illusive concept not only covers the development and refinement of a single isolated skill, but also the influence of previous experience and its influence on future learning.

Ever since Thorndike and Woodworth's (1901) formal introduction of the concept, transfer has been a topic touched on by numerous researchers. Some work in the field includes Singley (1989) and, most recently, Taatgen (2013).

Besides transfer scientists have also been interested in peoples' capacity to learn and reason at different ages. Inhelder and Piaget (1958) proposed a multi-stage model which has been supported by the results of many decision making tasks (Siegler 1976, 1981). We decided to use these stages in modeling Goldstein and Gigerenzer's (1996) Take the Best heuristic in the form of a decision making task.

Both Piaget's Balance Beam Task (BBT) and Gigerenzer's Take The Best heuristic (TTB) have been a focus of cognitive modeling (Van Rijn 2003, Nellen 2003). By using Taatgen's ACTransfer modeling framework (Taatgen, 2013) we can create models for both BBT and TTB and then analyze the transfer between them.

We chose these two tasks because they both produce distinct behaviors depending on the developmental stage of the subject. Specifically, the way in which multiple dimensions are handled changes depending on which stage the subject is currently in.

If task-specific skills are guided by task-general strategies we should see some of the general strategies being acquired in one task transfer to another. We explore this hypothesis by training on either the BBT and TTB task to see whether prior training on one seemingly unrelated task improves performance on the other. We can analyze both the transfer between tasks, but also the transfer between stages within a task.

At present, our work is free of empirical data and based solely on the concepts and parameters of ACTransfer. Hence, what we present here are theory-based predictions free of parameter tuning to empirical data.

Balance Beam Task

The balance beam task has been a focus of study for many years. This task was originally developed by Piaget (1958) and extensively studied by Siegler (1976, 1981), among many others. The task involves a subject (typically a child around age 11) being shown a balance beam with a certain number of equally massive weights on each side at certain specific distances from the fulcrum, with the beam being held from tipping in either direction. The subject is then asked which direction the balance beam will fall (or if it will stay balanced) upon release. After the subject submits their answer the beam is then allowed to tilt, showing the correct answer.

The methods used by children performing this task were said by Inhelder and Piaget to reflect developmental stages, and not experience-based changes in strategies. However, our immediate interest in this task is not with its history in developmental psychology but as a demonstration of general transfer with the ACTransfer theory. The 4 stages are:

Stage 1 Subject only considers a single dimension (in this case usually the number of weights on each side), the subject then makes a decision based on only that factor.

Stage 2 Subject considers various factors until one is found to support an individual side or all known factors are considered and therefore must balance.

Stage 3 Subject considers all factors and chooses the side with the most number of supporting factors (guesses if it's a tie).

Stage 4 The value of each factor is added up (or multiplied if they have been taught the torque rule), this is the final stage. Multiplication is usually not considered, unless explicitly taught, because addition so often yields the correct answer. This stage is not in the model.

Take The Best

Take the Best tasks (Gigerenzer, et al . 1996, 1999) involve a subject being given two options along with facts about each, and then choosing the best option. An example of this is determining which city is larger, given a list of facts about each. Take the Best can be used as a time-pressured algorithm that looks at the most important factors first. In time-pressured situations it is important to make a decision as quickly as possible, but there may not necessarily have a hard time limit. To make this model comparable to the balance beam task no time-pressure is considered. The stages in this task are similar to that in the balance beam task, except that in this task any number of dimensions can be considered (in this model we restrict it to three) as opposed to the balance beam's built-in limit of two. As well as the difference in the number of possible dimensions, Take the Best only has three stages (opposed to balance beam's four stages), these stages are:

Stage 1 Subject only considers first dimension and answers based on that.

Stage 2 Subject considers a sequence of new dimensions until a fact which supports a single option is found. If this evidence is not found and the list of known dimensions is exhausted than the subject will answer that all options are equal.

Stage 3 Subject considers all known dimensions and adds together how many dimensions support each option. Once all dimensions are analyzed the number of facts supporting each side are compared, if equal it is answered that all options are equal.

Previous Models

Models have been made for both the Balance Beam Task and Take the Best, each emphasizing their own features. The closest models to this one are Van Rijn's (2003) Balance Beam ACT-R model and Nellen's (2003) Take the Best ACT-R model (because ACTransfer is built on top of ACT-R). The Balance Beam Task has been focused on by mostly psychology and cognitive science while Take the Best is used in artificial intelligence for decision making in limited information, time-pressured environments.

ACTransfer

ACTransfer (Taatgen, 2013) is a cognitive modeling framework built on top of ACT-R, a widely used framework in the field of computational cognitive modeling (Anderson, 2007). The purpose of ACTransfer is to model and measure transfer between tasks. For example, the classically used cognitive experiment which tested typists learning to use different text editors, and assessed how learning one task influences how quickly one learns another, can be modeled very well in this framework. ACTransfer is able to model this phenomenon by breaking up traditional production rules in to more basic generalized elements called Primitive Information Processing Elements (PRIMs). PRIMs come in three different varieties, which all act upon slots in ACT-R's buffers. One kind copies information between slots, another kind compares two slots for equality, and the third writes to the slots from declarative memory (Taatgen, 2013).

As each combination of PRIMs is used those connections become stronger and faster, thus improving performance on other tasks that share the same combination of PRIMs. Not always do tasks that share PRIMs have surface similarities. ACTransfer has shown unexpected connections between several tasks (Taatgen, 2013). However, it is fair to say that the exploration of transfer between the TTB and BBT tasks represents a minor milestone in the application of the theory to predict general transfer.

ACTransfer models can be run to show the estimated time to complete certain tasks. These estimations can then be used to calculate the transfer between tasks with the equation developed by Katona (1940). In basic terms, the percent transfer is the improvement on the task by trial *n* after training on a different task, divided by the improvement on the task by trial *n* after training on the same task. If Pl(1) is performance on that task on trial 1, and Pl(n) is performance on trial *n* after performing that task on trials 1 through (n-1), and Pt(n) is performance on that task on trial *n* after performing a different task on trials 1 through (n-1): $\%Transfer = ((Pl(1) - Pt(n)) / (Pl(1) - Pl(n))) * 100$.

Transfer should be expected between any models that share similar production rules. The tasks analyzed here mostly share actions such as memory retrieval (translating data from the world into something usable by the model), buffer comparison (deciding if two values are equal), and response (an active response given by the model to choose an option). These rules become more complex as in later stages as illustrated by Figure 1 below.

Some similarities were unavoidable between the tasks, but others were focused on by the modeler with the logic that a single human would perform all of these tasks in a relatively similar way. Nonetheless, the transfer between models should be most representative of the similarities between the tasks.

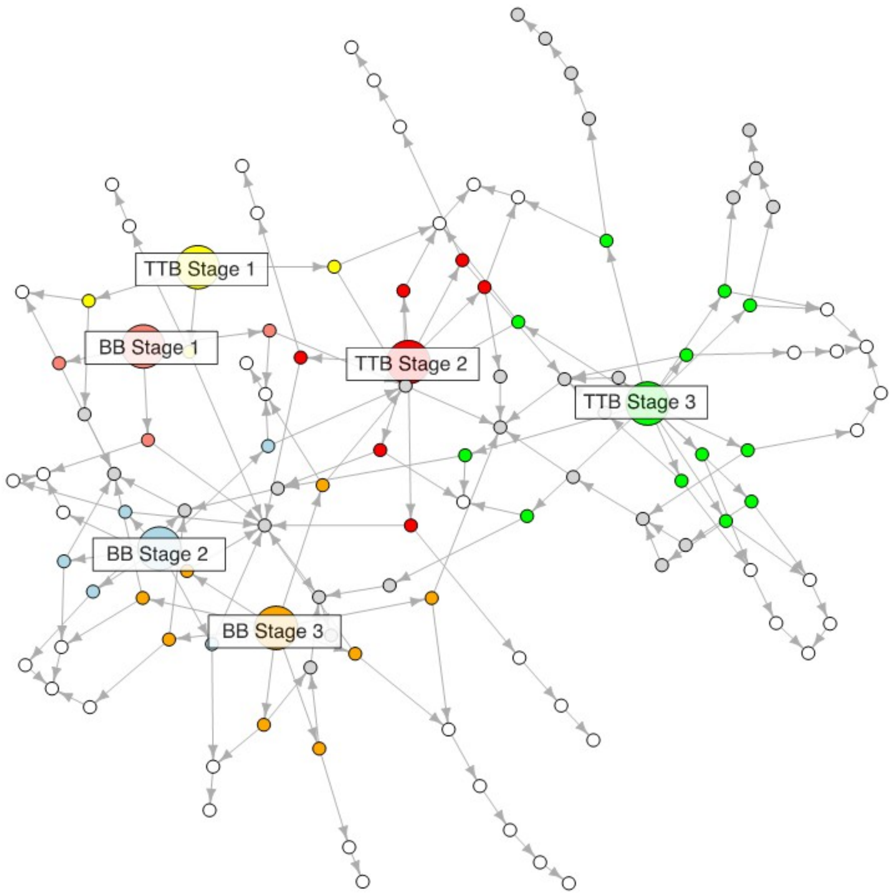

Figure 1: A graph of the various PRIMs (small circles) connecting the various models (large circles). (BB means Balance Beam Task)

The Model

The model created for this paper examines stages 1 through 3 of the Balance Beam Task and stages 1 through 3 of the Take the Best task using the ACTransfer framework. While stage 3 of both tasks are theoretically the same, they are modeled differently. Because the balance beam task is only comparing two dimensions, no sense of counting or number comparison is used, only congruency between the last remembered dimension and the current dimension. The Balance Beam models also do not contain an internal list of properties (unlike the Take the Best models) because they only look at two dimensions. It is plausible that the Balance Beam models in general do not need any sort of list. A list may be unnecessary because all that's needed for their heuristics is just a memory of whether they retrieved a property yet. This model does not aim to include any of the nuanced between stage properties like no-feedback transitions, rather it only looks at transfer. What we hope to see is transfer between the models, showing that learning one of the previous stages helps in learning the next stage and/or the other task.

Balance Beam Task Implementation

The balance beam task is implemented with three distinct stages. The first stage, which models the first psychological stage mentioned earlier, interacts with the simulated environment to observe the first property (semantically this is weight, but the model has no concept of that). The model then uses its declarative memory to retrieve the fact relevant to that observation, which it answers upon. Every time any of the models observes the environment they also perform this fact retrieval, the logic behind it being that an observation by itself is not usable knowledge until parsed.

The second stage of the balance beam model starts the same as the first, except it does not always answer immediately. If the first property does not explicitly support the left or right side falling (that the beam will balance) then the environment is queried again for a fact. When this second fact is retrieved (it is recognized as the second by virtue of its working memory being full) it is used to give the answer (because the first fact must necessarily have been 'balance').

			Tested on BALANCE BEAM TASK			Tested on TAKE THE BEST		
			Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
No Training		1st Trial (sec)	3.60	3.70	5.80	3.80	5.20	17.20
100 Trials of training on BALANCE BEAM TASK	Stage 1	101st Trial (sec)	1.40	2.10	4.00	2.80		
		Difference (sec)	2.20	1.60	1.80	1.00		
		Transfer	**100.00%**	**69.57%**	**48.65%**	**43.48%**		
	Stage 2	101st Trial (sec)		1.40	3.20		4.00	
		Difference (sec)		2.30	2.60		1.20	
		Transfer		**100.00%**	**70.27%**		**30.77%**	
	Stage 3	101st Trial (sec)			2.10			15.30
		Difference (sec)			3.70			1.90
		Transfer			**100.00%**			**14.84%**
100 Trials of training on TAKE THE BEST	Stage 1	101st Trial (sec)	2.50			1.50	3.80	16.30
		Difference (sec)	1.10			2.30	1.40	0.90
		Transfer	**50.00%**			**100.00%**	**35.90%**	**7.03%**
	Stage 2	101st Trial (sec)		3.00			1.30	13.40
		Difference (sec)		0.70			3.90	3.80
		Transfer		**30.43%**			**100.00%**	**29.69%**
	Stage 3	101st Trial (sec)			4.80			4.40
		Difference (sec)			1.00			12.80
		Transfer			**27.03%**			**100.00%**

Table 2: Results of training and testing the model within and between tasks.

The third stage of the balance beam model works the same way as the second stage, except that it always observes both dimensions and does not immediately answer based on the second observation. Instead the third stage uses the conditions of its production rules to determine if the two dimensions of evidence are congruous (in which case it answers the side they suggest), if one dimension suggests 'balance' (in which case it chooses the other dimension), or if the evidence is incongruous (in which case it guesses the answer).

Take The Best Implementation

For Take the Best, the first stage is similar to the first stage of the Balance Beam model. The largest difference, for both this stage and the rest of the Take the Best stages, being that the model observes a specific, predetermined, property in the environment. This is different than the balance beam in that the model actually specifies which fact it is observing (rather than just "the next fact"), after the first stage the model must use declarative and working memory to determine the next dimension.

The second stage acts with the same basic principle as the balance beam's second stage, except instead of considering only two dimensions it continuously observes new dimensions until either evidence is found to support one choice over the other or its list of known dimensions is exhausted. This stage could hypothetically consider any number of dimensions, unlike stages 2 and 3 of the Balance Beam model.

The third stage of the Take the Best model is more involved than any of the other stages. Like stage 3 of the Balance Beam model it considers all known dimensions before deciding, but instead of being able to just store a single dimension in working memory it must count how many pieces of evidence it observes in support for each option. After the list of dimensions is exhausted it compares the two numbers in working memory to determine which side has more evidence. In this model it can only count up to three, but with more facts in working memory it could work on any number of dimensions.

Results

Using Katona's transfer equation stated earlier with an *n* of 101 we found the results shown in Table 2.

The leftmost column denotes which task the model was trained on for the first 100 trials. The topmost row denotes which task the model was tested on for the 101st trial, except for the "no training" row in which the task was tested without any prior training (for the purpose of getting a baseline). The "Difference" rows show the difference in time between the model attempting the task with no prior training (the first row) and its attempt at the task after 100 trials of training (the task trained on is specified in the left-most column). The "Transfer" rows show the percent transfer derived from Katona's equation. For the stages testing is done on that stage only (e.g. Stage 3 Balance Beam Task training is independent of training any other stages, it simply means a completely fresh model was run on stage 3 for 100 trials before a 101st testing trial on whatever condition was specified)

Analysis

In the Balance Beam Task, the 70% transfer between stage 1 and stage 2 indicates that learning stage 1 does indeed help in learning stage 2, similar results are seen between stage 2 and 3 with a 70% transfer. The transfer between stages 1 and 3 is only 49% though, showing that the intermediate stage 2 clearly assists the transition.

Take the Best shows similar results, although not as strongly. 36% transfer between stage 1 and 2 show that they are related, but not as closely as the balance beam stages. 30% transfer between stage 2 and 3 show similar existent, yet weak results. The mere 7% transfer between stages 1 and 3 indicate that virtually no transfer occurs between the two, showing again that stage 2 is very important to transition to the final stage.

Horizontal transfer between tasks is not symmetrical, showing that going from Take the Best to Balance Beam carries, in general, higher transfer. This suggests that take Take the Best is a more complex task, which was expected given the increased number of dimensions. Looking at each stage individually, stage 1 carries the highest transfer with 43% going from the Balance Beam Task (BBT) to Take the Best (TTB) and 50% going from BBT to TTB, again this is reasonable given the simplicity and shared heuristics of the two stages. Stage 2 is the second most closely associated, with 30% transfer in either direction. Stage 3 shows the largest discrepancy in directionality with 15% going from BBT to TTB and 27% going from TTB to BBT. These values for stage 3 could be explained by the addition of counting and comparing values in the Take the Best model.

An interesting result of this model is that there is more transfer between Take the Best and Balance Beam at equivalent stages than there is between Take the Best at any particular stage and the next stage up. Notably, this same effect does not appear going from the Balance Beam Task to Take the Best. This conclusion seems to support the idea of a unified stage of development in one sense. This conclusion can also be explained as Take the Best being a task that has stages which are particularly difficult to advance through and the Balance Beam Task being generally easier.

One of the biggest problems with this paper's model is that some stages go through radically different paths of production rules depending on what kind of trial they are dealing with. For example, stage 2 in both the tasks can either be very quick, requiring a look at only the first dimension, or take a longer time, requiring looking at multiple dimensions if the first dimension does not indicate a specific option to choose. Because of this discrepancy the model will encounter certain levels of difficulty much less frequently than others, so the training on these rarer types of trials would be much less developed.

Conclusion

These models show how important intermediate developmental stages are to the development of decision making. Expecting a child to jump from stage 1 to stage 3, skipping stage 2, in any of these tasks would be unreasonable given the increased ease of going through the intermediate stage. Transfer between the tasks showed that going from a more complicated task to a less complicated one had the highest transferability, but initially learning the less complex task would help a child learn the more complicated one, especially at earlier developmental stages.

ACTransfer has proven itself to be a powerful tool in exploring learning and transfer that can lead to some surprising results. This framework could be used to support ideas such as Bavelier's (2003) notion of video games improving certain cognitive functions. It would be no surprise to see a task as complex as a video game to activate a plethora of PRIMs which are also used in a number of other tasks.

This framework could also push forward educational entertainment. Perhaps instead of directly teaching children the contents of traditional lesson plans we could use tasks similar to activities they willingly partake in during their free time. Having them learn these skills, with low-level connections to the desired skill that is trying to be taught, we could hide away rigorous practice in tasks that students enjoy, possibly even without their knowledge.

We see ACTransfer as ushering in a new age of modeling which could lead the way forward to discovering a whole array of connections between tasks that we never knew existed. Once these connections are discovered there lies endless possibilities for skill training and learning in general.

Acknowledgments

A special thank you to Prof. Niels Taatgen and the whole Cognitive Modeling lab at Groningen for a wonderful and productive semester. Thank you to Prof. Wayne Gray for supporting my research and assistance editing.

References

ACT-R v6.0. Retrieved March 15, 2013, from "ACT-R 6.0" Web site: http://act-r.psy.cmu.edu/actr6/.

ACTransfer v0.45. Retrieved March 15, 2013, from "Transfer: ACTransfer" Web site: http://www.ai.rug.nl/~niels/ACTransfer.html.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature,* 423(6939), 534-537.

Inhelder, B., and Piaget, J. (1958). *The Growth of Logical Thinking : From Childhood to Adolescence.* Basic Books, Inc.

Gigerenzer, G. & Goldstein, D. G. (1996). "Reasoning the fast and frugal way: Models of bounded rationality". *Psychological Review*, 103, 650-669.

Czerlinski, J., Goldstein, D. G., & Gigerenzer, G. (1999). "How good are simple heuristics?" In Gigerenzer, G., Todd, P. M. & the ABC Group, *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Katona, G. (1940). Organizing and memorizing: studies in the psychology of learning and teaching.

Nellen, S. (2003). The use of the "take-the-best" heuristic under different conditions, modeled with ACT-R. *In Proceedings of the fifth international conference on cognitive modeling (pp. 171-176).*

Rijn, H. v., Someren, M. v., & Maas, H. v.d. (2003). Modeling developmental transitions on the balance scale task. *Cognitive Science, 27,* 227-257.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology* 8:481–520.

Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development* 46(2):1-74.

Singley, Mark K. (1989) *The transfer of cognitive skill*. Harvard University Press, No. 9.

Taatgen, N.A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120(3),* 439-471.

Taatgen, N.A. (2013). The gap between architecture and model: strategies for executive control. *AAAI Technical Report FS-13-03* (pp. 95-101). Palo Alto, CA: AAAI Press.

Taatgen, N.A. (2013). Diminishing return in transfer: A PRIM model of the Frensch (1991) arithmetic experiment. Proceedings of the 12th international conference on cognitive modeling (pp. 29-34). Ottawa: Cartelon University.

Thorndike, E. L. and Woodworth, R. S. (1901) "The influence of improvement in one mental function upon the efficiency of other functions"*, Psychological Review 8*