ACT-R is *almost* a Model of Primate Task Learning: Experiments in Modelling Transitive Inference

Mark A. Wood (cspmaw@cs.bath.ac.uk)
Jonathan C. S. Leong (jleong@fas.harvard.edu)
Joanna J. Bryson (jjb@cs.bath.ac.uk)
University of Bath; Department of Computer Science
Bath, England BA2 7AY; United Kingdom

Abstract

We present a model of transitive inference (TI) using ACT-R which strengthens the hypothesis that TI is not dependent on underlying sequential ordering of stimuli, but rather on the learning of productions. We nevertheless find a weakness in the ACT-R sub-symbolic learning system and suggest improvements.

Introduction

The last decade has shown an increasing body of work indicating that ACT-R (Anderson and Lebiere, 1998) can be used to model human learning in an impressive variety of tasks. However, ACT-R has been slow to gain acceptance in mainstream experimental psychology as a useful model, possibly because it does not seem a very good correlate to the physical learning systems we find in the brain.

In previous and concurrent work, we have been exploring another model of task learning which also seems at first blush artificial and not particularly parsimonious, but has also shown an impressively tight fit to human and animal experimental data otherwise unaccounted for. This is the Harris (1988) production-rule-stack model of one of the main testbeds of task learning in the animal literature, the transitive inference task. We developed the two-tier model (Bryson, 2001; Bryson and Leong, 2004), which accounts for all of the Harris model's data, while extending the model to account for both learning and failing to learn this task (a common outcome in live subjects). We've also found potential neurological correlates for the two-tier model.

The two-tier model hypothesises two learning systems: one for connecting perceptual contexts to actions, and another for prioritising which of those perceptual contexts to attend to if more than one are present simultaneously. We realised that this model had similar aspects to ACT-R, which also has two learning systems, one symbolic and one statistical. We therefore decided to apply ACT-R to the transitive inference learning task. Our results show that ACT-R is far better than the standard TI models at accounting for the particular (and somewhat controversial) data set that prompted the Harris (1988) model, and for some individuals provides a better model than Harris (1988), though for others it cannot. Our results lead us to believe that the two-tier model is the best existing model of transitive inference, although ACT-R is close enough that it is probably fixable. ACT-R demonstrates one significant simplification over the two-tier hypothesis, and has one important difference from real mammalian task learning.

Transitive Inference

Transitive inference (TI) formally refers to the process of reasoning whereby one infers that if, for some quality, $A > B$ and $B > C$, then $A > C$. In some domains, such as integers or heights, this property will hold for any A, B or C, though for others it does not (see Wright, 2001, for a recent discussion). TI is classically described as an example of concrete operational thought (Piaget, 1954). That is, children become capable of doing TI when they become capable of mentally performing the physical manipulations they would use to determine the correct answer, a stage they reach at approximately the age of seven. In the case of TI, this manipulation involves ordering the objects into a sequence using the rules $A > B$ and $B > C$, and then observing the relative location of A and C.

Since the 1970's, however, apparent TI has been demonstrated in much younger children (Bryant and Trabasso, 1971) and a variety of animals, from monkeys (McGonigle and Chalmers, 1977) to pigeons (Fersen et al., 1991) — not normally ascribed with concrete operational abilities. The behaviour of choosing A from AC without training after having previously been trained to select A from AB and B from BC is consequently sometimes referred to as "transitive *performance*", and whether it implies sequential ordering at all is now an open issue.

The main motivation for *not* considering TI in animals to be based on a sequential structure is a dataset due to McGonigle and Chalmers (1977), which they have subsequently replicated both with monkeys and children. This data set concerns what happens if subjects demonstrating TI are asked to select between *three* items rather than two. Some individuals show significant, systematic degradation in performance, which cannot be explained by a sequential model. Some researchers have dismissed the triad data as resulting from confusion in the subjects due to the extra item. These criticisms were dealt with in a replication by McGonigle and Chalmers (1992) which provided the main data set used in this paper and by Harris and McGonigle (1994). The fact that the systematicity of the degradation has now been successfully accounted for further validates this data set. This data concerns monkeys trained on 4 adjacent pairs drawn from a 5 item sequence, AB, BC, CD, DE.

The Models

Due to space constraints we will not review the more traditional, sequence-based or simple-associative models of TI, but see further (Wynne, 1998; Bryson and Leong, 2004). These models cannot account for the triad data set.

Harris and McGonigle

Harris (1988) showed that both pair and triad TI performance could be accounted for if we assume that monkeys learn a production rule stack. A *production rule* is a basic AI representation which connects a stimulus to a response. A *stack* is a prioritised list. In the Harris model, each monkey learns one rule per possible stimulus, or up to 5 rules in total. One of two actions is associated with each rule, either *select* or *avoid*. If a subject applies the rule $A \rightarrow s(A)$ (see A implies select A), then it will simply pick up A, regardless of whether other items are present. However, if a subject applies the rule $A \rightarrow a(A)$ it will pick up anything *but A*. If more than one other item is present, the subject is at chance for which object it will pick up. If more than one rule could apply, then whichever rule is higher in the stack (has higher *priority*) will be applied.

Although Harris' hypothesis may seem obscure, it shows a remarkable match to the data. If one assumes that rules are limited to the case that the action refers to the object attended to, then only 16 of the 1920 ($10 \times 8 \times 6 \times 4$) possible stacks of four rules operate correctly on all training pairs (Harris and McGonigle, 1994). All 16 of these stacks also correctly perform TI on *all* pairs automatically, thus already accounting for one of the mysteries of transitive performance.

The degradation some subjects display on the triad tests is a consequence of the random aspect of the *avoid* rules. In fact, triads can be used to discriminate which rule stack an individual subject has learnt. For example, a stack that consists entirely of selects ($s(A), s(B), s(C) \ldots$) will never make any errors. One that starts with $a(E)$ will be at chance between the other two options whenever E is present in a triad.

Table 1 shows all of the possible discriminable stacks as identified by Harris and McGonigle (1994). Because only the highest-priority applicable rule fires and there are always at least two applicable rules (since there are at least two stimuli), there is no way to discriminate the two lowest-priority rules using triad performance. These stacks therefore only reflect the top three rules of the stacks.

Table 1: Enumeration of Harris and McGonigle Stacks

#	Rule Depth 1	2	3	#	Rule Depth 1	2	3
1	s(A)	s(B)	s(C)	5	a(E)	s(A)	s(B)
2	s(A)	s(B)	a(E)	6	a(E)	s(A)	a(D)
3	s(A)	a(E)	s(B)	7	a(E)	a(D)	s(A)
4	s(A)	a(E)	a(D)	8	a(E)	a(D)	a(C)

Table 2: Primate TI training régime (Chalmers and McGonigle, 1984; McGonigle and Chalmers, 1992)

P1	Each pair in order (DE, CD, BC, AB) repeated until 9 of 10 most recent trials are correct Reject if requires over 200 trials total
P2a	4 of each pair in order. Criteria: 32 consecutive trials correct. Reject if requires over 200 trials total
P2b	2 of each pair in order. Criteria: 16 consecutive trials correct. Reject if requires over 200 trials total
P2c	1 of each pair in order. Criteria: 30 consecutive trials correct. No rejection criteria
P3	1 of each pair randomly ordered Criteria: 24 consecutive trials correct Reject if requires over 200 trials total
T	Pair and triad testing

The Two-Tier Model

A successfully trained two-tier model creates a replication of the production-rule-stack model (Bryson, 2001). However, the two-tier model is dynamic, and as such gives us insight into why animals have trouble learning the initial pairs for the TI task, the sorts of mistakes they may make, and the impact of training régimes. The first tier of the two-tier model is a single-vector neural network (NN) which learns the prioritisation of the stimuli. The second tier is a set of small two-item vectors which each learn to associate an action with one of the stimuli. The learning rule for the NNs is a slight simplification of standard delta learning (Widrow and Hoff, Jr., 1960).

Simulations using the two-tier model show artificial subjects successfully learning the training data only about 25% of the time when training pairs are presented in a random order. However, switching to the training régime applied by McGonigle and Chalmers (1992) shown in Table 2, which is standard for primates, the success rate increases to about 75%, which is comparable to live subjects (Bryson and Leong, 2004).

Further, the sorts of errors made by artificial subjects failing to learn are consistent with those shown by live subjects — they tend to confuse the middle pairs. Analysis of the networks shows that this is nearly always a consequence of misprioritising the rules representing the end pairs. An agent can guarantee it always selects A in the pair AB (the only pair A appears in) by learning $a(B)$, and there is a great inclination to learn about middle rules because these are the ones that have the most data (and the most confusing data, since B is sometimes rewarded but sometimes penalised.) However, there are no successful stacks which do not have one end point or the other at the highest priority (see Table 1). The training régime greatly increases the probability of learning correct prioritisation. For further details see Bryson and Leong (2004).

ACT-R

As for the above models, ACT-R also learns production rules, but any number of these rules may have their preconditions for firing satisfied at any given time. In this case, ACT-R's conflict-resolution system selects the rule with highest *utility* value

Rule utilities are changed by ACT-R's sub-symbolic processing system It is possible to attach *success* or *failure* tags to productions and when such a rule is fired, ACT-R backtracks to discover which rules fired previously and increments or decrements their utilities respectively More precisely, the utility of a rule is given by:

$$U = PG - C + \epsilon(s) \quad (1)$$

where G is the *goal value*, C is the *expected cost*, $\epsilon(s)$ is the *expected gain noise* and P is the *expected probability of success*:

$$P = \frac{\text{Successes}}{\text{Successes} + \text{Failures}} \quad (2)$$

In our experiments, rather than make arbitrary changes to ACT-R's many available parameters in an attempt to best fit the data, we have used mostly defaults. The most notable exception to this is that we set the initial Failure count to 1 which, along with ACT-R's default setting of 1 for Successes[1], gives an initial probability of success of 0 5. This change was also made by Belavkin and Ritter (2003) in their Dancing Mouse model. To maximise the number of successful agents, we also tried a range of values for s (which affects the variance of the noise function), finally deciding upon $s = 1$.

One trial consists of two or three items displayed onscreen which the agent encodes into its goal buffer. The goal state is then changed, enabling it to make decisions about which item to pick (see below) Once an item has been picked, either a reward or no reward is displayed appropriately, the agent notes its success or failure respectively and the next trial begins

We have tested two different approaches to solving the TI problem in ACT-R. In the first, the *select* and *avoid* rules for each item are independent, concurrent candidates for execution. For three displayed items, this corresponds to six conflicting rules that have their preconditions satisfied Henceforth we refer to this approach as *ACT-R-1*

In the second, the agent must *focus* on a displayed item before either *select*ing or *avoid*ing it, as in the two-tier model This results in an extra stage of conflict-resolution for the agent With three options there are at first three candidate *focus* rules whose actions alter the agent's goal state. This, in turn, satisfies two further rules; *select* and *avoid* for the focus-item. We call this approach *ACT-R-2*

Results

As with the two-tier model and live subjects, our ACT-R model produces both agents that successfully learn the task and agents that do not 44 of the 100 agents tested with ACT-R-1 successfully passed the training régime. This compares to 35% of those using ACT-R-2, or 75% of those using the two-tier system We examine these groups separately

Successful Agents

The stack model proposed by Harris and McGonigle (1994) attempts to fit the McGonigle and Chalmers (1977) triad data to any of the eight discriminable correct stacks (Table 1). In contrast, the ACT-R agents learn only two possible solutions

There are two rules which are always successful for all pairs: s(A) and a(E) This means that, provided these rules are discovered by the agent, their utilities will begin to converge to maximal, given by:

$$\begin{aligned}\lim_{t\to\infty} U &= \lim_{t\to\infty} (PG - C) \quad (3)\\ &= \lim_{t\to\infty} \left(\frac{\text{Successes}}{\text{Successes} + \text{Failures}}\right) G - C\\ &= \lim_{t\to\infty} \left(\frac{1}{1 + \frac{\text{Failures}}{\text{Successes}}}\right) G - C\\ &= G - C\end{aligned}$$

where t is the number of trials, and G and C remain constant at ACT-R default values throughout. Therefore any successful ACT-R agent will have these two rules at highest priority. This eliminates half of the Harris and McGonigle stacks, leaving 3, 4, 5 and 6 (Table 1)

In addition, because the top-two rule utilities are converging to the same value, it becomes essentially arbitrary (in fact, governed by the expected gain noise) whether s(A) or a(E) occupies the top stack position for any given choice. In other words, the ACT-R agents do not learn a totally ordered stack, but effectively a pair of stacks. The two possible pairs are:

Hybrid Stack 1 (HS1): s(A)a(E)s(B) & a(E)s(A)s(B)
Hybrid Stack 2 (HS2): s(A)a(E)a(D) & a(E)s(A)a(D)

Table 3 shows each triad with the expected percentage of trials in which each item in that triad is chosen These probabilities are the same for both Hybrid Stacks, except for the triad BCD. In this case, the format is HS1 / HS2 A 75%/25% split occurs when both A and E are present in the triad. We assume that half the time s(A) has top priority and is thus selected. Otherwise, a(E) has top priority, giving an even chance of A or the other item being selected

Taking into account the noise added to the system, this model well describes the behaviour of many of the ACT-R agents In 45% of cases, one of the Hybrid Stacks fitted the agent's distribution better than any of the individual stacks, and a further 47% were best fitted by Stack 5; a contributor to HS1

Failed Agents

Despite these encouraging results, a majority of the ACT-R agents failed the training régime (Table 2), which is not true of the monkeys (albeit there were only seven

[1] The default is Failures = 0 $\Rightarrow P_{initial} = 1$

Table 3: Projected percentage choice distributions for Hybrid Stack 1 / 2

Triad	A	B	C	D	E
ABC	100	0	0	-	-
BCD	-	100 / 50	0 / 50	0 / 0	-
BDE	-	50	-	50	0
CDE	-	-	50	50	0
BCE	-	50	50	-	0
ABD	100	0	-	0	-
ACD	100	-	0	0	-
ADE	75	-	-	25	0
ABE	75	25	-	-	0
ACE	75	-	25	-	0
Mean	52.5	22.5 / 17.5	12.5 / 17.5	12.5	0

test subjects) or children (Chalmers and McGonigle, 1984) Virtually all of the agents which failed did so at stage P2a (Table 2), typically having seen less than 300 training pairs in total. There are two exceptions for both ACT-R models which failed at P3

To best understand why agents fail, we examine each training pair and determine what causes agents to pick the wrong item:

AB Since s(A) almost always has a high utility, errors on this pair tend to be caused by interference from s(B), whose utility is driven up by its success on pair BC. Ironically, then, it is agents who are too successful too soon who fail because of this pair, training stage P2a having the most stringent pass criteria.

BC C is picked when s(C) is too high relative to s(B) or, less frequently, a(C). Occasionally a(B) adds to this interference but, due to the success of s(A), rarely attains a high enough utility

CD The symmetric case of BC. Here a(C)/a(D) interference is the chief cause of error (see Figure 1)

DE As for AB, if a(D) is discovered early to be a good rule, it interferes with a(E) causing small but significant errors in P2a

For some agents (around 25%), failure is a result of a combination of the above interferences If many of the interfering rules are interdependent (eg s(B), s(C), a(C), a(D)) then this can lead to a more even distribution of errors across all training pairs Conversely, if two sets of independent rules (eg. s(A), a(B), a(D), a(E)) are interfering, often two training pairs are consistently incorrect, with little or no error on the other two

As Tables 4 and 5 show, agents confuse the middle pairs far more often than the end pairs (see also **Analysis**). This, in turn, is most often the result of s(C) and a(C), both of which perform incorrectly for one of the middle pairs We have restricted these data to the last 200 trials carried out by the failed agent. This focuses on the specific phase of training at which the agent failed and removes the noisiest choices, made during P1.

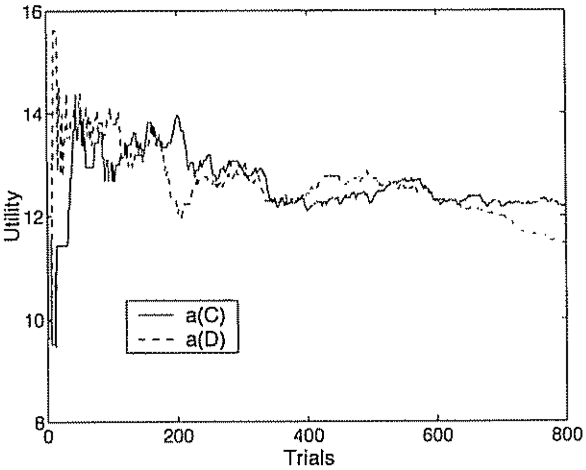

Figure 1: a(C) and a(D) fight for control of the pair CD

Table 4: Aggregate percentage error on each pair

Group	AB	BC	CD	DE	Mean
ACT-R-1	6	25	30	7	17
ACT-R-2	6	43	39	9	24

Table 5: Percentage distribution of failed agents

	Modal Error Pair			
Group	AB	BC	CD	DE
ACT-R-1	2	39	54	5
ACT-R-2	0	42	50	8

Analysis

For ease of statistical comparison, we applied the χ^2 test in the same way as Harris and McGonigle (1994): by excluding item E, which (usually) has an expected value of 0.

For three of the five individual test subjects for whom McGonigle and Chalmers (1977) triadic data is available, one of the hybrid stacks fits better than any of the eight others (Table 6). As explained in the **Successful Agents** section above, and in contrast with the Harris and McGonigle stacks, the hybrid stacks do not represent a total ordering Thus it would seem that neither do some monkeys form a total ordering, and their choices cannot be perfectly modelled by a simple production-rule system For Bump and Brown, however, our model is rejected ($p < 0.01$), suggesting that other monkeys do come up with a total ordering, which cannot be well modelled in ACT-R

The two-tier model does support both models, although its admittedly simplistic learning rule tends to favour the total ordering These results suggest that

Table 6: Comparison of individual triadic choice data (1977) with both Hybrid and Harris' Stack models

	A	B	C	D	E	χ^2	p(O)
Bill	55	17	20	8	0	-	-
HS2	52.5	17.5	17.5	12.5	0	2.1	n.s.
S 4	60	15	15	10	0	2.8	n.s.
Blue	55	25	14	6	0	-	-
HS1	52.5	22.5	12.5	12.5	0	4.0	n.s.
S 3	60	20	10	10	0	4.9	n.s.
Bump	53	34	8	4	1	-	-
HS1	52.5	22.5	12.5	12.5	0	13.3	< 0.01
S 2	60	30	5	5	0	3.4	n.s.
Brown	36	29	24	11	0	-	-
HS2	52.5	17.5	17.5	12.5	0	15.3	< 0.01
S 7	35	25	25	15	0	1.8	n.s.
Roger	51	26	6	17	0	-	-
HS1	52.5	22.5	12.5	12.5	0	5.6	n.s.
S 5	45	25	15	15	0	6.5	< 0.1

an improved priority-learning rule for either the two-tier model or ACT-R could result in a highly accurate model of TI and possibly task learning in general.

There were just five monkeys who passed criteria and so were included in the triad phase of the 1977 experiment. These five only represented three of the eight stacks in Table 1. This may render the grouped data unrepresentative, but our ACT-R model still displays a better correlation than Harris and McGonigle (1994) of α-choices, as shown in Table 7, where α represents the *correct* choice in a given triad (see also Table 8).

Table 7: Correlation of α-choices to group data

Group	r	p
ACT-R-1	0.688	$p < 0.05$
ACT-R-2	0.692	$p < 0.05$
Hybrid Stack Model	0.616	$p < 0.1$
H & M Stack Model	0.634	$p < 0.05$

Upon closer examination of the choices made for each triad, we see ACT-R closely matching the monkey data for those triads which do not contain the item E (Table 8). For those that do, ACT-R makes more mistakes, suggesting that the monkeys do not have a(E) at as high a priority. This might reflect a primate bias against having identical priorities for rules.

There may be a good reason for favouring priorities over utilities for ordering rules. For example, there is no circumstance in which an ACT-R agent can reach a stable enough solution to reduce error to zero. Suppose such a situation was attained. Then every decision made would result in success and thus increase the utility of the executed rule. Eventually, these rules (of which there must be at least three to produce a correct stack) would converge upon the same value (see Equation 3 above). But since no three rules in a correct stack are independent for all triads, they will start to interfere with each other, causing error to be re-introduced.

This phenomenon is best demonstrated by examining the errors of agents who did not take part in structured training, but were presented with the training pairs in a random order. Here, as is usual, the s(A) and a(E) rules have high, convergent utilities (these rules *are* independent for all training pairs). Then the utility of one (or both) of the other successful rules, s(B) and a(D), will also start to converge. This results in errors made on the *end pairs* since s(B) interferes with s(A) for AB (Figure 2) and a(D) interferes with a(E) for DE. Neither s(C) nor a(C) can attain a high utility, because they will perform incorrectly on one of the middle pairs (BC and CD). The final result is that the middle pairs are chosen consistently correctly, whereas the end pairs have small errors (typically around 15%). This certainly seems somewhat biologically implausible, and contradicts the *End-Anchor Effect* (Bryant and Trabasso, 1971; Wynne, 1998).

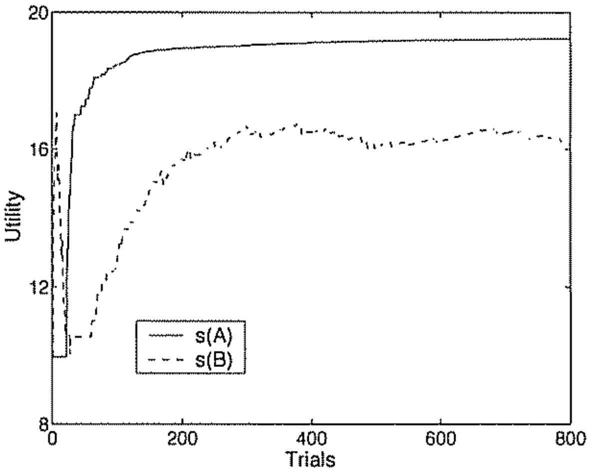

Figure 2: Interference with s(A) prevents the utility of s(B) reaching above a certain level

Conclusions and Further Work

The ACT-R models lack in their ability to represent stable, totally ordered stacks, which some real subjects appear to form. On the other hand, the Harris and McGonigle (1994) stacks lack the flexibility to represent more dynamic solutions to the TI problem. For this reason we conclude that the two-tier model is the best existing model of TI. On the other hand, the fact that there is no significant difference between ACT-R-1 (where no initial item focus is required) and ACT-R-2 (where this focus is required) implies that the two-tier model can be simplified to allow arbitrary numbers of stimulus action pairings, as is the default case in ACT-R.

Table 8: Percentage of items selected - triadic analysis

Triad	Monkeys			ACT-R-1			ACT-R-2			Hybrid Stack Model			H & M Stack Model		
$\alpha\beta\gamma$	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ
ABC	80	18	2	83	17	0	86	14	0	100	0	0	94	6	0
BCD	70	26	4	70	29	1	72	26	2	75	25	0	75	25	0
BDE	66	34	0	59	35	6	56	37	6	50	50	0	63	38	0
CDE	62	38	0	49	41	10	48	44	7	50	50	0	56	44	0
BCE	78	22	0	58	42	0	58	42	0	50	50	0	63	38	0
ABD	80	20	0	79	21	0	80	20	0	100	0	0	88	13	0
ACD	86	12	2	90	9	0	91	8	0	100	0	0	88	13	0
ADE	86	14	0	72	24	4	71	26	4	75	25	0	75	25	0
ABE	88	12	0	68	32	0	70	30	0	75	25	0	75	25	0
ACE	80	20	0	76	24	0	74	25	0	75	25	0	75	25	0
Means	78	22	1	70	28	2	71	27	2	75	25	0	75	25	0

There are several obvious next steps First, as stated in the **Introduction**, the learning rules for priorities in both the two-tier model and ACT-R need improvement, though in different ways. We will be focusing on improving the two-tier model, but would be happy to see or support ACT-R being modified to reflect these results. Also, we suggest two possible improvements to the ACT-R model Allowing ACT-R to compile its own system of rules from a minimal starting set (Anderson and Lebiere, 1998) may provide a more natural solution, although interpreting the underlying decision processes would be more difficult Starting with a high initial noise would allow the agents to always discover and benefit from the most successful rules, while rapidly reducing this noise level (in conjunction with the *entropy of success* (Belavkin and Ritter, 2003)) would be necessary in order to obtain a stable enough solution to pass stage P2a of the training régime

The other obvious next step would be to collect and analyse more triad testing results across a larger number of primates For our purposes, it would be useful to have triad testing on subjects who fail TI training as well as those who succeed We are investigating collaborations in this area

The greatest significance of this work is that it gives further evidence for a non-sequence-based representation underlying the TI task and further supports the utility of the McGonigle and Chalmers (1977) triad data set.

References

Anderson, J. R and Lebiere, C (1998) *The Atomic Components of Thought* Mahwah, NJ: Erlbaum.

Belavkin, R V and Ritter, F E (2003). The use of entropy for analysis and control of cognitive models. In Detje, F , Dörner, D , and Schaub, H , editors, *Proceedings of the Fifth International Conference on Cognitive Modelling*, pp 21–26 Universitäts-Verlag Bamberg

Bryant, P E and Trabasso, T (1971) Transitive inferences and memory in young children. *Nature*, 232:456–458

Bryson, J J (2001). *Intelligence by Design: Principles of Modularity and Coordination for Engineering Complex Adaptive Agents.* PhD thesis, MIT, Department of EECS, Ch. 9 AI Technical Report 2001-003

Bryson, J J and Leong, J C S (2004) A two-tiered model of primate transitive performance: Separate memory systems for rule-stimulus association pairs and their prioritisations In preparation

Chalmers, M and McGonigle, B O (1984). Are children any more logical than monkeys on the five term series problem? *Journal of Experimental Child Psychology*, 37:355–377

Fersen, L , Wynne, C D L , Delius, J., and Staddon, J E R (1991). Transitive inference formation in pigeons *Journal of Experimental Psychology: Animal Behavior Processes*, 17:334–341

Harris, M R (1988) *Computational Modelling of Transitive Inference: a Micro Analysis of a Simple Form of Reasoning* PhD thesis, University of Edinburgh

Harris, M. R. and McGonigle, B O (1994). A model of transitive choice *The Quarterly Journal of Experimental Psychology*, 47B(3):319–348.

McGonigle, B. O. and Chalmers, M (1977) Are monkeys logical? *Nature*, 267:694–696

McGonigle, B. O and Chalmers, M (1992). Monkeys are rational! *The Quarterly Journal of Experimental Psychology*, 45B(3):189–228.

Piaget, J. (1954) *The Construction of Reality in the Child* Basic Books, New York.

Widrow, B and Hoff, Jr , M. E (1960). Adaptive switching circuits *IRE WESCON Convention Record*, 4:96–104.

Wright, B C (2001) Reconceptualizing the transitive inference ability: A framework for existing and future research. *Developmental Review*, 21(4):375–422

Wynne, C D L. (1998) A minimal model of transitive inference. In Wynne, C D L and Staddon, J E R , editors, *Models of Action*, pages 269–307 Lawrence Erlbaum Associates, Mahwah, NJ