INSTANCE-BASED COGNITIVE MODELS OF DECISION-MAKING

By Cleotilde Gonzalez and Christian Lebiere*

I. INTRODUCTION

'Cognitive architectures' are computer algorithms designed to model human behavior and to function in a way similar to the workings of the human mind. The breadth of cognitive architectures is one of their primary strengths. Rather than serving as special-purpose models engineered specifically for individual tasks, cognitive architectures provide general computational mechanisms and constraints that are applicable to the development of models for all kinds of tasks.

ACT-R is a widely researched cognitive architecture that accounts for hundreds of empirical results obtained in the field of experimental psychology (Anderson and Lebiere, 1998). ACT-R is a hybrid architecture of cognition that combines a production system (to capture the sequential, symbolic structure of cognition) with a subsymbolic, statistical layer (to capture the adaptive nature of cognition). A goal of ACT-R researchers is to investigate the overall integration of cognition by building models designed to explain how all the components of the mind work together (Anderson, 2002).

Although cognitive architectures like ACT-R can offer flexibility and precision in human-like behavior representation, they have rarely been used to study economic decision making. A reason for this state of affairs is that ACT-R has mistakenly been

*With the support of the Advanced Decision Architectures Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory (DAAD19-01-2-0009) and the Office of Naval Research (N00014-01-10677).

conceptualized as a rule-based static theory that does not provide the flexibility necessary for uncertain decision situations, like economic settings. This chapter will demonstrate the potential of ACT-R to model economic decision making.

Economic decision making should be modeled as a learning process, involving more than calculation of expected values, accounting for human cognitive limitations and abilities, and allowing for flexibility in transfer of knowledge. This chapter summarizes evidence of successful ACT-R modeling of decision making processes of this kind. Several examples of ACT-R decision making models show that the same architecture can be used in a variety of tasks including dynamic control tasks, backgammon players and simple 2 x 2 gamers like in the Prisoner's Dilemma. We argue that for economic decision making settings as well as for many other tasks in which learning and decision making occur in unison, instance-based decision making is the most plausible learning mechanism (Gonzalez, Lerch and Lebiere, 2003). Other researchers have also theorized that instance-based decision making is a general mechanism used for all types of decision making under uncertainty (Gilboa and Schmeidler, 1995). All the models reported in this chapter have successfully used this instance-based approach in ACT-R, concluding that ACT-R can provide an integrated account of the psychology of decision making.

The rest of this chapter is organized as follows: Section 2 presents an introduction to the ACT-R cognitive architecture, its knowledge representation structures and memory and learning mechanisms. Section 3 summarizes the instance-based decision making approach from the psychology and economics perspectives. Section 4 presents a compilation of results from instance-based ACT-R models in individual decision making tasks and in 2 x 2 economic games. Section 5 further demonstrates the use of ACT-R instance-based decision making models in complex and more real-world tasks. Section 6 concludes.

II. ACT-R COGNITIVE ARCHITECTURE

ACT-R 5.0 (Figure 1) is a modular, neurally plausible architecture structured as a set of localized modules (e.g., long-term memory, visual, motor) that interact through limited-capacity buffers connected to a central production, pattern-matching module.[1]

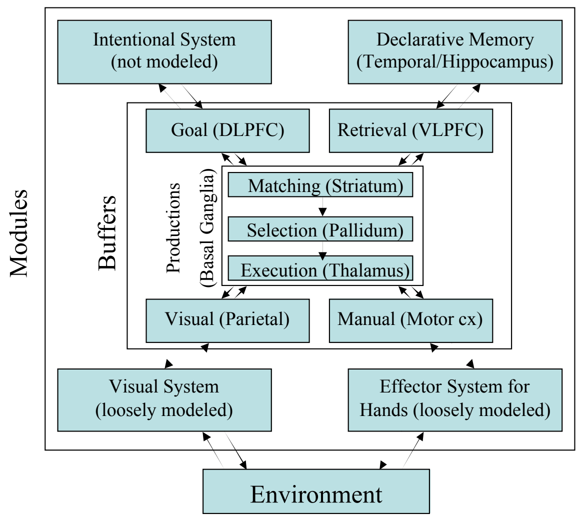

FIGURE 1. ACT-R 5.0 ARCHITECTURE

Although the perceptual/motor modules effectively constrain performance in a number of tasks and provide a principled model of interaction with the environment, they are not our focus in this chapter. Rather, we chose to explore high-level decision making,

---

[1]Current hypotheses regarding the neural location of the various modules and buffers are indicated in parentheses.

for which declarative memory and the central production system are the modules of foremost importance.

ACT-R incorporates a symbolic system in which declarative knowledge and procedural knowledge interact in discrete cycles. Declarative structures called ‘chunks’ are used to store factual knowledge in the declarative memory. Chunks encode knowledge as structured, schema-like configurations of labeled slots. ‘Productions’ are modular, condition-action rules that encode procedural memory by representing potential actions to be taken when certain conditions are met. ACT-R also incorporates a subsymbolic system in which continuously varying quantities are processed simultaneously to produce many of the graded characteristics of human cognition. These subsymbolic quantities participate in neural-like activation processes that determine the speed and success with which decision makers access chunks in the declarative memory and resolve conflicts among productions. Finally, ACT-R also incorporates a set of learning processes that can lead to the creation of new symbolic knowledge structures and the modification of the subsymbolic quantities associated with those structures.

The subsymbolic activation processes believed to be implicated in instance-based decision making make a memory chunk available for retrieval to the degree that the similarity between a past experience and a current context (as defined by a current goal) indicates the usefulness of the chunk at that particular moment. Retrieving a chunk results in its immediate reinforcement (due to its frequency of use) through ACT-R’s base-level activation learning mechanism. Activation (1) reflects, in Bayesian terms, the log posterior odds that a chunk is relevant in a particular situation. The activation A_i of a chunk i is computed as the sum of its ‘base-level activation’ B_i plus its ‘context activation’:

$$A_i = B_i + \sum_j W_j S_{ji} \tag{1}$$

In determining the context activation, W_j designates the attentional weight given the context element j. An element j is in context if it is part of the current goal chunk (i.e., the value of one of the goal slots). S_{ji} stands for the strength of association between an element j and a chunk i. ACT-R assumes that there is a limited source activation capacity shared equally by each goal element. Source activation capacity is typically assigned a value of 1. Thus if there are n source elements for the current goal, each element receives a source activation of 1/n (Anderson, Reder, and Lebiere, 1996). The 'associative strength' S_{ji} between an activation source j and a chunk i is a measure of how often chunk i was needed (i.e., was retrieved in a production) when source j was in the context. Associative strengths provide an estimate of the log likelihood ratio, which measures how much the presence of a cue j in a goal slot increases the probability that a particular chunk i must be retrieved to instantiate a production.

The base-level activation of a chunk (2) is determined by using an architectural mechanism incorporating the past history of use of a chunk i:

$$B_i = \ln \sum_{j=1}^{n} t_j^{-d} \approx \ln \frac{nL^{-d}}{1-d} \tag{2}$$

In the above formula, t_j stands for the time elapsed since the j^{th} reference to chunk i, d represents the memory decay rate, and L denotes the lifetime of a chunk (i.e., the time since its creation).

Researchers in psychology have demonstrated that both, forgetting and learning are characterized mathematically by a power function (Rubin and Wenzel, 1996; Newell and Rosenbloom, 1981). For example, plotting the logarithm of the time to perform a task against the logarithm of the trial number always yields a straight line (Newell and Rosenbloom, 1981). Anderson and Schooler (Anderson and Schooler, 1991) have shown that the base-level learning equation produces both the Power Law of Forgetting and the Power Law of Learning. Strengths of association are determined by using a similar mechanism that records the statistics of co-occurrence between sources and retrieved chunks (see Anderson and Lebiere, 1998 for more detail).

When retrieving a chunk (3) to instantiate a production, ACT-R selects the chunk with the highest activation A_i. However, some stochasticity is introduced within the system by adding Gaussian noise of mean 0 and standard deviation σ to the activation A_i of each chunk. To be retrieved, the activation of a chunk needs to reach a fixed retrieval threshold τ that limits the accessibility of declarative elements. If the Gaussian noise is approximated with a sigmoid distribution, the probability P of chunk i being retrieved by a production is determined as follows:

$$P = \frac{1}{1 + \exp^{-\frac{A_i - \tau}{s}}} \tag{3}$$

where $s = \sqrt{3}\sigma/\pi$. The activation of a chunk i is directly related to the latency of its retrieval by a production p. Formally, retrieval time T_{ip} is an exponentially decreasing function of the chunk's activation A_i:

$$T_{ip} = Fe^{-A_i} \tag{4}$$

where F is a time scaling factor. In addition to determining the latencies for chunk retrieval (provided by equation (4), the total time required for selecting and applying a production is determined by executing the actions of a production's action part, with a value of 50 ms typically assumed for elementary internal actions. External actions, such as pressing a key, usually have a longer latency that is determined by the ACT-R/Pm perceptual-motor module (Byrne and Anderson, 1998).

Instead of only retrieving chunks that perfectly match the production conditions, ACT-R's 'partial-matching' mechanism (5) can retrieve whichever chunk matches the condition to the greatest degree, according to a similarity function. Specifically, the chunk with the highest match score is retrieved, where match score M_{ip} is a function of the activation of chunk i in production p and its degree of mismatch to the desired values:

$$M_{ip} = A_i - MP\sum_{v,d}(1 - Sim(v,d)) \qquad (5)$$

In the above formula *MP* is a mismatch penalty constant, while *Sim(v,d)* stands for the similarity between the desired value *v* held in the goal and the actual value *d* held in the retrieved chunk, and permits the representation of continuous quantities. Thus even if no chunk in memory perfectly matches a current context, a common occurrence given an infinite number of continuous values, the chunk holding the closest value can be retrieved if its match score, after subtracting the mismatch between values from its activation, remains higher than the retrieval threshold (and the match scores of competing chunks).

One shortcoming of partial matching is that, although it generalizes the matching process to handle continuous quantities, it can only return a value already present in some chunk. Lebiere (1999) proposed ‘blending’, which is a generalization of the retrieval mechanism and allows the retrieval and averaging of values from multiple chunks rather than a single one, thereby enabling the generation of continuous values. This powerful type of interpolation has proved useful for a range of paradigms of implicit learning (Gonzalez *et al.*, 2003; Wallach and Lebiere, 2003). Specifically, the value obtained by a blended retrieval is determined as follows:

$$V = Min\sum_{i} P_i(1 - Sim(V,V_i))^2 \qquad (6)$$

where P_i is the probability of retrieving chunk i and V_i is the value held by that chunk. Thus blending can be viewed as the process of returning the value that best satisfies conflicting pieces of knowledge stored in memory.

Blending also represents a generalization of well-known AI techniques. Neural networks have a similar ability to learn in their connection weights a number of training patterns and produce an output that reflects the constraints of the entire training set rather than any specific pattern. The Bayes Optimal Classifier produces the most likely outcome weighted over all hypotheses (ACT-R chunks), rather than simply the most likely hypothesis (most active chunk). Linear weighted regression is an instance-based machine learning algorithm that produces the answer that minimizes the squared error between a fitted function and a set of data points, with each data point being weighted by its distance to the query point. The blending mechanism combines attributes of all these techniques.

In summary, the ACT-R cognitive architecture incorporates a set of mechanisms that can be used to develop models of learning and performance. The assumptions described in this section are not arbitrary; they are supported by years of work resulting in the accurate modeling of a broad range of results in experimental psychology.

III. INSTANCE-BASED DECISION MAKING

Making decisions based on instances basically means that courses of action are chosen through the use of accumulated experience. Observations in real-world, complex situations (e.g., military battle executions and firefighting) support the notion that under conditions involving stress, uncertainty, or task overload, people’s decision making is mostly experience-based (Klein, Orasanu, Calderwood and Zsambok, 1993; Pew and Mavor, 1998; Zsambok and Klein, 1997). Two theories, one from economics and the other from psychology, crystallize this form of decision making.

The economists Gilboa and Schmeidler (1995) proposed a theory called 'case-based decision theory' (CBDT) (Gilboa and Schmeidler, 1995) designed to explain decision making under conditions of uncertainty. Like expected-utility theory, CBDT derives a functional representation of preferences from a set of axioms about individual behavior. But, in contrast to expected-utility theory, CBDT posits that decision makers rely on their experience by choosing alternatives that have worked best in the past. Central to the theory is the concept that memory consists of a finite set of past instances or cases and that similarity to past cases is the *only* guide to decision making. Cases in CBDT are triplets (q, a, r) where q is the problem situation, a is the act (decision), and r is the consequence resulting from that act in the situation q. CBDT assumes that decision makers judge both the similarity of the problems they encounter and the desirability of the outcomes (i.e., the utility) as they are experiencing them. Gilboa and Schmeidler (1996) also emphasize that the similarity function is only a matter of subjective judgment and is not "objectively known" by the decision maker. Although CBDT can provide an axiomatic derivation of a similarity function, it does not explain how individuals select similar past cases for comparative purposes (Gilboa and Schmeidler, 1996). A similarity function and an accumulation of cases in the memory would not necessarily help a decision maker select the "right choice"—i.e., basing a decision solely on which case in the memory is most similar to a current situation may yield a poor outcome. Thus, when using past cases to make decisions individuals presumably take both the similarity of the problem to past cases and the past cases' utility (i.e., a measure of the desirability of outcomes) into account.

Working in the field of psychology, Gonzalez, Lerch and Lebiere (2003) proposed a theory called 'instance-based learning theory' (IBLT) in an effort to help explain decision making in complex, dynamic situations. This theory builds on memory theories of learning developed in psychology and on theories of decision-making under conditions of uncertainty (e.g., CBDT). Like CBDT, IBLT proposes that instances (or cases) accumulate in the memory as individuals make decisions, and incorporates the concepts of similarity and utility. But, in contrast to CBDT, IBLT describes a cognitively plausible decision-making process through which individuals acquire and update instances and through which learning occurs.

A large part of the difficulty associated with developing accurate models of human decision making stems from the fact that much of the knowledge gained and deployed by experts (and even minimally skilled novices) in decision-making situations seems to be implicit. Berry and Broadbent (1984) were the first researchers to explore the concept of implicit learning in process control, and they reported negative correlations between task performance and the ability to answer specific questions about a system's behavior (Berry and Broadbent, 1984; Broadbent, 1977). ACT-R provides a straightforward theory of the difference between knowledge and performance based on the acquisition and retrieval of instances. Each instance in the form of a declarative chunk is a piece of conscious knowledge; however, both the process of retrieving and applying that knowledge and the subsymbolic parameters that control that process (e.g., base-level activations, strengths of associations, and similarities between chunks) are consciously inaccessible and constitute the implicit knowledge of the system.

IBLT claims that every execution of a dynamic decision-making task results in the creation of instances. These instances are represented as chunks with slots containing the situation (a set of cues), the decision made, and the expected utility of that decision in that situation. IBLT proposes that initially (i.e., in the absence of accumulated knowledge relevant to a task) decision makers evaluate alternatives by using simple heuristics (e.g., random choice). However, as individuals acquire knowledge in the form of instances, they use these instances in subsequent executions of the task. The feedback process, another key characteristic of IBLT, updates the utility slot according to the outcome of decisions. Thus, decision makers confronted with similar situations while performing a task gradually abandon general heuristics in favor of improved instance-based decision-making processes.

The different mechanisms used to retrieve instances, evaluate alternatives, and apply feedback are central to IBLT. A similarity mechanism plays an integral role in individuals' recognition of decision-making situations. If the situation is relatively common, then the use of past experience might enable individuals to make accurate, relatively fast decisions. But if the similarity between a current situation and existent memory instances is low, then the use of a more general heuristic might be the most efficient method by which to select a course of action. IBLT also posits that, in addition to the analysis of situation similarity, a sequential evaluation of alternatives influences individuals' decisions. Thus the order in which alternatives are evaluated by decision makers acting in a complex environment while under time constraints is very important. This component of IBLT, called 'the necessity mechanism', is very important when individuals encounter environmental or cognitive constraints that limit information seeking. Finally, a feedback mechanism determines the amount of credit afforded to individual decisions given a particular outcome. This mechanism is particularly important in dynamic situations, in which multiple and interdependent decisions typically occur before the final outcome is known.

Despite the differences between CBDT and IBLT, they share some common principles, including the conceptualization of memory as a set of past instances or cases and of retrieval as a process based on some sort of similarity and utility metrics. However, IBLT is particularly well-suited for modeling decision making that is constrained by dynamics and environmental factors such as uncertainty, time constraints, and high workload. The next section reviews multiple examples of instance-based decision-making models developed in ACT-R research.

IV. ACT-R INSTANCE-BASED MODELS OF DECISION MAKING

This section summarizes a set of ACT-R models of decision-making tasks performed by either individuals or small teams. Researchers have developed these models in accordance with the instance-based approach and have validated them by using human data. We briefly summarize how accurately data generated through the use of each of these models describe actual human behavior.

The sugar factory task (Berry and Broadbent, 1984) is an instance-based ACT-R model. Sugar factory is a computer-simulated task in which participants are told to imagine that they are factory managers and can control the production of sugar *sp* by determining the number of workers *w* employed on each of a number of trials. Unbeknownst to the participants, the following equation governs the behavior of the system:

$$sp_t = 2 \times xw_2 - sp_{t-1} \qquad (7)$$

Sugar production is proportional to the number of workers employed, a concept that is intuitive enough, but is inversely related to the sugar production at the previous step, a relationship that is difficult and counterintuitive to infer. The value entered for the workers hired (w_t) can be varied in 12 discrete steps ($1 \leq w_t \leq 12$), while the sugar production sp_t changes discretely within the range $1 \leq sp_t \leq 12$. To allow for a more realistic interpretation of *w* as the number of workers and *sp* as tons of sugar, the actual computer simulation multiplies these values by 100 and 1000, respectively. If the result according to the equation is less than 1000, *sp* is simply set to 1000. Similarly, a result greater than 12000 always leads to an output of 12000 tons of sugar. Finally, in two-thirds of all trials a random value of ±1000 is added to the result derived from the equation above.

Dienes and Fahey (1995, 1998) developed two models of the sugar factory task using either rules or instances and found that the former model reproduced human behavior more closely than the latter. In addition, they found that subjects that displayed the best control performance of the system also exhibited the lowest amount of system knowledge, as determined by a post-task test.

Wallach and Lebiere developed an ACT-R instance-based learning model of the sugar factory and compared it to a models proposed by Dienes and Fahey (Dienes and Fahey, 1995, 1998; Wallach and Lebiere, 2003). The ACT-R model is quite simple, consisting of a single heuristic rule (taken from the Dienes and Fahey model) to bootstrap the system and another rule to retrieve past instances. Nonetheless, in comparison with the models of Dienes and Fahey it provides at least as good of a fit to human data without making any unwarranted assumptions. The ACT-R model also explains lowest amount of system knowledge in best performers: the model's knowledge of the system consists only of instances rather than any general, abstract understanding of the system's dynamics.

Because the sugar factory task has only a few discrete states (a single input control variable and a single output variable), Wallach and Lebiere (2003) tested the generality of the ACT-R modeling approach by applying it to Broadbent's transportation task (Broadbent, 1977; Broadbent and Aston, 1978). Participants performing that task can adjust two continuous input variables to try to achieve target values on two continuously-varying output variables. Like the sugar factory equation, the equations underlying the dynamics of the transportation task system contain not only straightforward relationships between input and output variables but also a counterintuitive negative cross-relationship. The proposed ACT-R model challenges the views of Berry and others by substantiating the ability of an instance-based learning model—representing pairs of encountered input-output values without explicitly encoding structural knowledge about causal relationships between variables—to successfully control the task (Berry and Broadbent, 1987). The model makes use of the ACT-R blending mechanism that retrieves the value or values that best satisfy the constraints expressed by an entire set of chunks, with each chunk weighted by its probability of retrieval.

Figure 2 shows that the model's ability to control the system is quite comparable to that of the test subjects', with r^2 of .73. Control performance was measured by the number of trials necessary to achieve the respective target value pairs. The average number of errors, defined as an increase in the distance to the required target values from trial n to trial $n+1$, was within the empirically observed range.

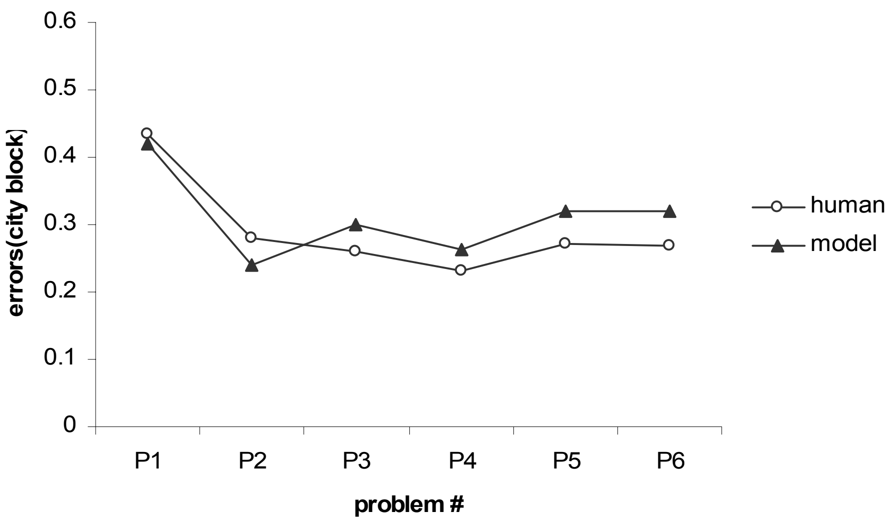

FIGURE 2: AVERAGE NUMBER OF ERRORS AS A FUNCTION OF PROBLEMS

The ACT-R model of the transportation task has an intriguing characteristic: A set of instances representing each subject's exploration phase is used instead of a general heuristic rule to initialize the instance-based model. Thus each model run constitutes an individualized version of the general instance-based model, adjusted to the knowledge of each individual subject.

Researchers also have applied the instance-based models of learning by individual decision makers to decision making by two-person teams. The mechanisms used to model decision making in multi-player game settings are the same as those used to model individual decision making. Instead of using theories specially developed for the task, such as game theory, we have built models based on the same general-purpose mechanisms of the ACT-R cognitive architecture (e.g., learning and memory).

Researchers use game-like tasks to evaluate team decision making because the competitive aspect of game playing is a good tool by which to ensure maximal effort by subjects and to test the limits of the subjects' cognitive abilities. Because these multi-person adversarial games involve a finite, often small number of choices that are repeated for a certain number of iterations, they allow for instance-based learning to occur.

Instance-based decision making is largely dependent on an individual's ability to match current situational patterns with past situations and the associated decisions stored in memory. In team decision-making situations, memory consists of one's own instances and those of the opponent. Thus, the quality of one player's decisions often depends strongly upon her awareness of the opponent's instances and upon her ability to analyze them to infer her opponent's plans.

To identify the essence of team tasks, Lebiere and West (1999) studied human decision making in the classic game of paper-rock-scissors. This game embodies the essence of adversarial decision-making: It offers each player a finite set of options, has simple, well-defined zero-sum[2] rules that define the outcome of those options, and allows

---

[2]A zero-sum game is one in which every gain by one of the players has to be offset by an equivalent loss by another. It has been recently argued that many real-world situations can in fact be

the most direct expression of the dynamic character of move and countermove, as each player tries to anticipate the other's moves and preempt them. Rather than attempting to create a model specifically tailored for this particular task, in accordance with the architectural approach we re-used a model previously developed for the basic human skill of learning sequences of events. The ACT-R Sequence Learning Model learned sequences of stimuli by building instances encoding short pieces of sequences (Lebiere, Wallach and Taatgen, 1998). Although the model accumulates instances straightforwardly through experience, the procedural knowledge is quite trivial and consists basically of a pair of production rules that match the most recent move against instances in memory, retrieve the most active instance with its prediction of the next move, then lead to selection of the move that counters the predicted one.

Figure 3 presents a time course of model and subject performance for a number of sample runs. The model stored the opponent's moves as sequences of different lengths (called 'lags'). For instance, if the model stored the opponent's most recent 2 moves, together with the current move, it had a sequence of length 2 and was termed a 'lag2' model. Figure 3 shows the mean score differential between the lag2 model and the lag1 model. Although the differential in score between the lag2 and lag1 models fluctuates, the long-term trend is clearly in favor of the more powerful lag2 model. Figure 3 also shows the mean score differential between human subjects and the lag1 model, and indicates that the lag2 model generally provides data very similar to that generated by human subjects. West and Lebiere (2001) present a more extensive analysis of the model and its sensitivity to a number of parameters, including length of stored sequence, impact of feedback, and various system parameters. Although paper-rock-scissors is a simple

---

characterized as non-zero-sum (Wright, 2001), but our approach is not dependent upon the zero-sum characteristic of situations and generalizes to non-zero-sum situations, as illustrated later in the chapter.

game, playing it well is by no means trivial. This generalized ACT-R model was entered in an international competition[3] and placed in the top tier, holding its own against specialized AI programs designed specifically for the game.

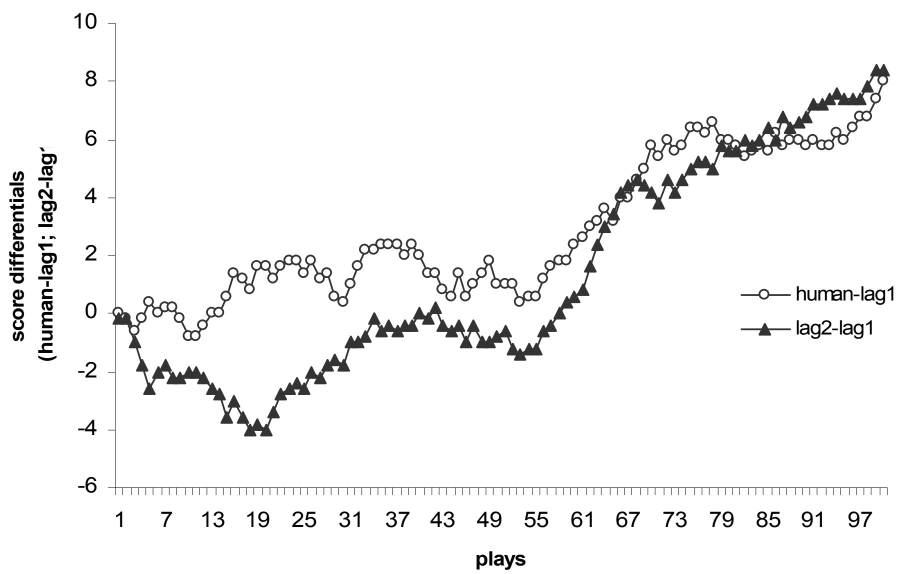

FIGURE 3: MEAN OF HUMAN AGAINST LAG1 MODEL AND MEAN OF LAG2 AGAINST LAG1 MODEL

The generality of the two-person game-playing model was validated in other two-person game situations. Rapoport *et al.* (1976) provide a wealth of data for a variety of 2x2 games. Among those games are some classic conundrums, such as the Prisoner's Dilemma (PD). PD brings the factors of cooperation and coordination into play, in that the players can achieve a better combined outcome by cooperating with each other rather

[3]Details of the competition are available at http://www.cs.ualberta.ca/~darse/rsbpc.html.

than by trying to maximize their own separate outcomes. Lebiere, Wallach and West (2000) present a model of the PD directly based upon the paper-rock-scissors model. They argue (as detailed above) that the chunks stored from experience in declarative memory contain the record of each trial, including one's own move, the other player's move, and the associated payoff. The PD decision is made in accordance with a pair of production rules that, given each possible action (i.e., cooperation or defection), retrieve the most likely outcome from memory and then select the one with the highest payoff. Aside from a slight variation in the model to reflect the new situation of varying payoffs in non-zero sum games, all parameters in the model remained unchanged. Again, the declarative knowledge simply represents a direct encoding of the player's experience, and the decision rule is a straightforward encoding of the rules of the situation. The model predictions again originate from the architectural learning occurring automatically at the subsymbolic level of the architecture.

Table 1A presents both the frequencies of the four possible outcomes for each pair of PD human subjects and the frequency average over the 10 pairs. The results are strongly bimodal. Six of the 10 pairs exhibited what can be described as cooperating behavior, choosing the cooperative outcome (B1B2) for two-thirds of the plays or more. Three of the 10 pairs exhibited non-cooperating behavior, choosing the defecting outcome (A1A2) more than half of the time. The percentage of choice of the non-symmetrical outcomes (i.e., A1B2 and A2B1) is fairly low, averaging 7% and 8%, respectively, except in pair #8, in which the B1A2 outcome was chosen more than half of the time. Overall, subjects selected the cooperative outcome more than half of the time (55%) and the mutually defecting outcome less than one-third of the time (30%).

TABLE 1: FREQUENCIES OF THE FOUR OUTCOMES IN THE PRISONER'S DILEMMA

(A) HUMAN

Subject Pair	A1A2	A1B2	B1A2	B1B2
1	1	1	1	97
2	7	1	1	92
3	14	1	2	83
4	04	5	5	86
5	21	4	3	72
6	24	5	5	66
7	54	12	7	27
8	34	2	52	11
9	58	25	5	12
10	83	9	4	3
Mean	30	7	8	55

(B) MODEL

Model Run	A1A2	A1B2	B1A2	B1B2
1	10	13	12	65
2	1	0	2	97
3	4	19	12	65
4	92	4	3	1
5	93	3	3	1
6	1	1	2	96
7	95	3	2	0
8	13	21	18	48
9	2	9	2	87
10	5	4	10	81
Mean	32	8	6	54

Two ACT-R models were run in pairs, interacting with each other for the same number of trials completed by the human subject pairs (10). Table 1B shows the corresponding frequencies for the model. Remarkably, the model matched not only the mean percentage of outcomes over the 10 pairs, but also the distribution of outcomes across pairs, with approximately the same number of pairs cooperating and defecting (and even a mixed outcome pair). Figure 4 indicates that the model also reproduced the time course of the human subjects' gradual shift in decisions from the original preponderance of defection toward more cooperation.

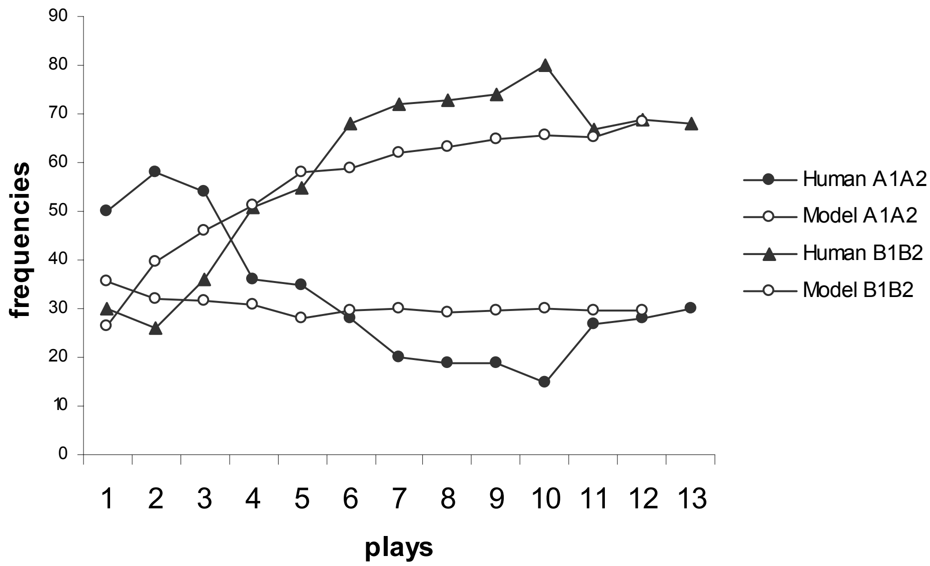

FIGURE 4: FREQUENCIES OF OUTCOMES OVER TIME IN THE PRISONER'S DILEMMA: HUMAN AND MODEL DATA

To further test the generality of the model, Lebiere *et al.* (2000) applied it to 10 other 2x2 games described by Rapoport *et al.* (1976). Despite the broad range of games, the only alteration made to the model when applied to the different games was to change the payoff matrix implemented in the task. The correlation between outcome percentages was 0.825, which supports the model's ability to predict the outcome of the decision-making process when applied to new situations that involve very different types of behaviors.

V. SCALING UP TO COMPLEXITY, UNCERTAINTY, AND DELAYED FEEDBACK

Although most of the situations confronted by the previously described models were not trivial and clearly captured some fundamental aspects of human decision-making, they shared a certain simplicity that is not reflective of real-world complexity. At each decision point, subjects (or the model) had a small number of discrete options (typically 2 or 3). The decision makers also received immediate feedback from their actions, which greatly aided in the learning process. Finally, although the decision-making process itself may have added a measure of uncertainty, the actual tasks were entirely deterministic. When studying human cognition it is essential to ascertain if the instance-based decision-making approach can indeed deal with more complex situations. Researchers have documented the use of this approach in complex tasks at the individual level (Gonzalez *et al.*, 2003).

The IBLT process has been implemented in the ACT-R architecture in the context of a dynamic task requiring resource allocation and scheduling (Gonzalez *et al., 2*003). This dynamic decision-making task, known as the 'Water Purification Plant' (WPP), involves uncertainty and feedback loops generated by the interrelationship of a user's decisions. WPP requires individuals to "purify" water via different treatment processes while acting under a deadline. They make these decisions while a simulation clock runs and must react as water arrives in different tanks and in unknown patterns.

This ACT-R model of IBLT relies heavily on the symbolic representations of ACT-R, but also reflects the subsymbolic processes described by the activation equation (and thus, the base-level learning equation), partial matching, and blending. Each instance has an activation value that depends on attention, base-level activation and learning, and other probabilities. Thus, rather than relying solely on past instances to guide decision making, IBLT incorporates many cognitive phenomena and mechanisms for decision making.

Researchers have used human data to validate this ACT-R model of IBLT. Gonzalez *et al.* (2003) articulate the step-by-step process by which WPP actuates the cognitive mechanisms proposed by IBLT. Figure 5 shows only the best fit of the model to human performance, with an r^2 of .90. This fit resulted from a process in which the ACT-R model evaluates alternatives one by one, selecting them randomly when first exposed to the task and according to their expected utility after acquiring knowledge in the form of instances. The model uses a large number of instances, even if these instances are only slightly similar to the current situation, rather than one instance that is exactly the same as the current situation. Eventually, the model learns to react at the right speed to the changes in the environment and not to use much of the feedback, but rather to learn from the interaction with the environment.

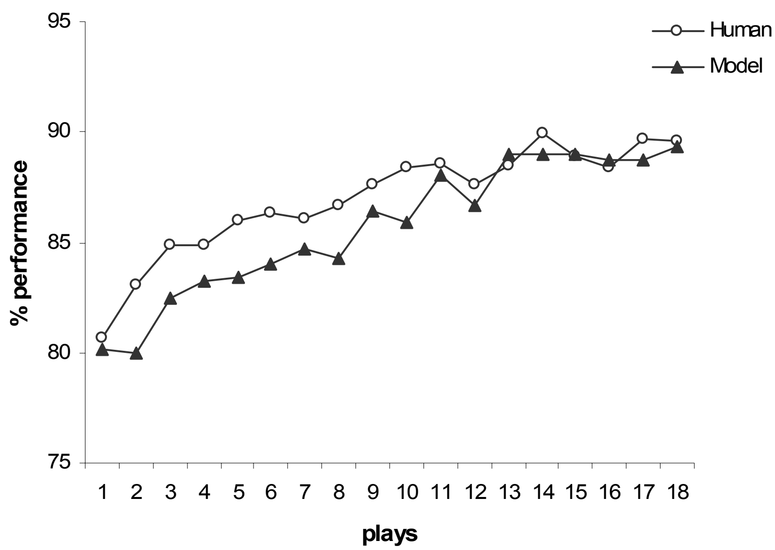

FIGURE 5. MODEL AND HUMAN DATA COMPARISON OF PERFORMANCE OVER THE COURSE OF 18 TRIALS

Although the ACT-R model of WPP addresses issues of decision making in complex environment, many decision making tasks in the real world are performed in a team. The instance-based approach also suits team decision making. The well-known game of backgammon presents a slightly complex team situation characterized by delayed feedback and a large degree of uncertainty in the task itself. Backgammon is a two-person board game involving substantial complexity. Because players can use up to 30 pieces, and place each piece in 1 of 24 possible positions, there exist an exponentially large number of combinations and up to hundreds of possible moves at each step. Furthermore, players receive no definite feedback until the game is won or lost after as many as over a hundred moves. Finally, the use of dice in backgammon introduces an element of uncertainty that greatly dilutes the effectiveness of the players' look-ahead searches.

Researchers have developed an ACT-R backgammon model that uses the same instance-based decision-making principles described above (Sanner, Anderson, Lebiere and Lovett, 2000). Essentially, the model represents each possible move by breaking it down into its fundamental features, such as capturing a piece or forming a block, and thereby keeps each memory chunk manageably small, resulting in a fundamental cognitive constraint. Each feature in memory accumulates its history of contributing to wins and losses, and the model yields an evaluation of a particular move by combining and analyzing the various features constituting that move. The model selects the most promising move available at each step. Because high performance requires high sensitivity to subtle positional changes, the representation includes the exact position of each piece on the board. However, since some specific feature positions are unlikely to be seen often enough to build accurate representations, the model generalizes across similar positions when necessary by employing the partial matching (and blending) component of the activation equation, which allows retrieval of closely matching chunks. This sort of similarity-based generalization essentially captures the effect of distributed representations in connectionist networks.

Starting with this general representation, the model gradually builds its declarative knowledge based on the experience gained by playing against a relatively strong opponent. A publicly available evaluation function (Tesauro, 1992) was used to play with the ACT-R backgammon model. The model was trained for 1000 games against the 'expert' opponent. Figure 6 displays the percentage of games won over the 1000 games for the ACT-R model and the opponent's model. The results indicate the model requires approximately 100 games to learn to play relatively well and almost matches the performance of its strong opponent by 1000 games (See Sanner et al., for further analyses).

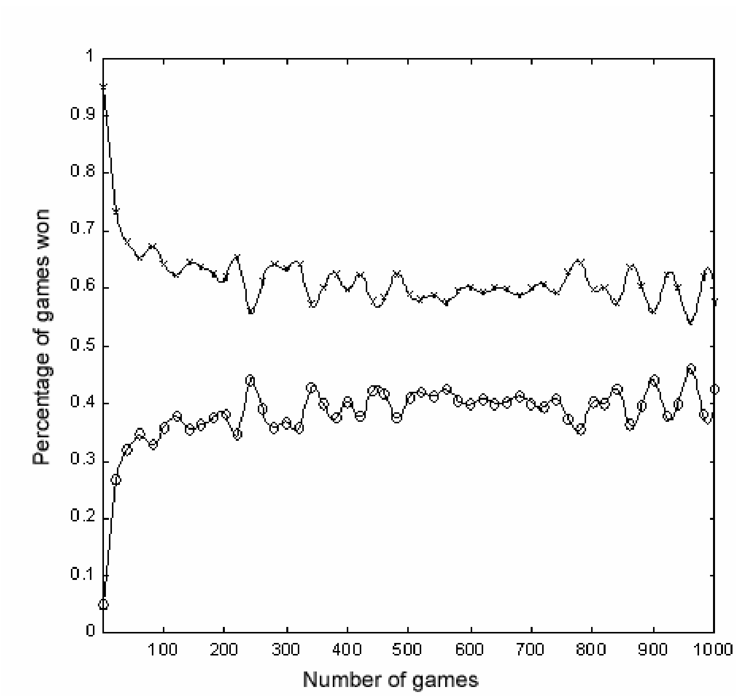

FIGURE 6: (A) PERCENTAGE OF WINS (O) AGAINST OPPONENT (X)

VI. CONCLUSION

This chapter describes ACT-R as a cognitive architecture that facilitates the development of decision-making models. In particular, we have summarized the instance-based approach to modeling decision making.

The models of decision making summarized in this chapter are based on two basic principles: a) the storage of instances in declarative memory and b) production rules that generate decisions by comparing the current situation to previous instances and selecting the most promising course of action by reviewing past experiences. Most importantly, the presented models are psychologically plausible—i.e., they are based upon a validated cognitive architecture and they can learn on a human-like scale of experience without requiring the excessive engineering of representational features. In general, results from multiple studies indicate the instance-based models that use the ACT-R learning mechanisms show sensitivity to strategic factors not built into the models but instead learned from experience.

Despite the success stories, cognitive modeling has often used to merely incorporate data fitting exercises that tweak the model *post hoc* to reproduce the subject data. Roberts and Pashler (2001) have raised this point forcefully, and researchers in the cognitive modeling community take this constraint seriously (Pew and Gluck, 2001; Roberts and Pashler, 2000). Cognitive modeling can provide a more useful insight into decision-making processes if the models make *a priori* predictions when applied to new situations. As more models are built for an increasing range of tasks, both the power and predictiveness of the cognitive architectures will increase.

Models of economic decision-making settings can benefit from a cognitive architecture like ACT-R. Not only is ACT-R a powerful computational architecture that combines most of the mechanisms required for economic settings, but compared to other computational approaches, ACT-R also provides a more realistic characterization of the flexibility and adaptability of human behavior. ACT-R's learning mechanisms supported by psychological research can effectively explain and represent transfer of knowledge.

Carnegie Mellon University, US.
Contact email: **coty@cmu.edu**

REFERENCES

Anderson, J. R. (2002), "Spanning seven orders of magnitude: A challenge for cognitive modeling", *Cognitive Science,* 26, 85-112.

Anderson, J. R. and Lebiere, C. (1998), *The atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. Reder, L. M., and Lebiere, C. (1996), "Working memory: Activation limitations on retrieval. *Cognitive Psychology,* 30, 221-256.

Anderson, J. R. and Schooler, L. J. (1991), "Reflections of the environment in memory", *Psychological Science,* 2, 396-408.

Berry, D. C. and Broadbent, D. E. (1987) "The combination of explicit and implicit learning processes in task control", *Psychological Research,* 49, 7-15.

Berry, D. E. and Broadbent, D. E. (1984), "On the Relationship between Task Performance and Associated Verbalized Knowledge", *The Quarterly Journal of Experimental Psychology,* 36A, 209-231.

Broadbent, D. E. (1977), "Levels, hierarchies, and the locus of control", *Quarterly Journal of Experimental Psychology,* 29, 181-201.

Broadbent, D. E. and Aston, B. (1978), "Human control of a simulated economic system", *Ergonomics,* 21, 1035-1043.

Byrne, M. D. and Anderson, J. R. (1998), "Perception and action", In J. R. Anderson and C. Lebiere (Eds.), *The atomic components of thought* (pp. 167-200). Mahwah: Lea.

Dienes, Z. and Fahey, R. (1995), "Role of specific instances in controlling a dynamic system", *Journal of Experimental Psychology: Learning, Memory and Cognition,* 21, 848-862.

Dienes, Z. and Fahey, R. (1998), "The role of implicit memory in controlling a dynamic system", *Quarterly Journal of Experimental Psychology: Human Experimental Psychology,* 51A, 593-614.

Gilboa, I. and Schmeidler, D. (1995), "Case-Based Decision Theory", *The Quarterly Journal of Economics,* 110, 605-639.

Gilboa, I. and Schmeidler, D. (1996), "Case-based knowledge and induction", *unpublished manuscript*.

Gonzalez, C., Lerch, J. F. and Lebiere, C. (2003), "Instance-based learning in dynamic decision making", *Cognitive Science,* 27, 591-635.

Klein, G., Orasanu, J., Calderwood, R. and Zsambok, C. E. (Eds.). (1993), *Decision Making in Action: Models and Methods*, Norwood, New Jersey: Ablex Publishing Corporation.

Lebiere, C., Wallach, D. and Taatgen, N. (1998), *Implicit and explicit learning in ACT-R,* Paper presented at the Second European Conference on Cognitive Modeling, Groningen, The Netherlands.

Newell, A. and Rosenbloom, P. S. (1981), "Mechanisms of skill acquisition and the law of practice", In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Earlbaum.

Pew, R. W. and Gluck, K. A. (2001), *Overview of the agent-based modeling and behavior representation (AMBR) model comparison project,* Paper presented at the 10th Annual CGF-BR Conference.

Pew, R. W. and Mavor, A. S. (1998), *Modeling Human and Organizational Behavior*, Washington: National City Press.

Roberts, R. D. and Pashler, H. (2000), "How persuasive is a good fit? A comment on theory description of retention", *Psychological Review,* 103, 734-760.

Rubin, D. C. and Wenzel, A. E. (1996), "One hundred years of forgetting: A quantitative description of retention", *Psychological Review,* 103, 734-760.

Sanner, S., Anderson, J. R., Lebiere, C. and Lovett, M. C. (2000), *Achieving efficient and cognitively plausible learning in Backgammon,* Paper presented at the Seventeenth International Conference on Machine Learning, San Francisco.

Tesauro, G. (1992), *Temporal Difference Learning of Backgammon Strategy*, IBM Thomas J. Watson Research Labs Technical Report, pp. 1-6

Wallach, D. and Lebiere, C. (2003), "Conscious and unconscious knowledge: Mapping to the symbolic and subsymbolic levels of a hybrid architecture", In L. Jimenez (Ed.), *Attention and implicit learning.* Netherlands: Amsterdam: John Benjamins Publishing Company.

Wright, R. (2001), *Nonzero: The logic of human destiny*, New York: Vintage Books.

Zsambok, C. E. and Klein, G. (Eds.). (1997), *Naturalistic Decision Making*, Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.