CHAPTER 10

Micro-Process Models of Decision Making

Jerome R. Busemeyer and Joseph G. Johnson

1. Introduction

Computational models are like the new kids in town for the field of decision making. This field is largely dominated by axiomatic utility theories (Bell, Raiffa, & Tversky, 1998; Luce, 2000) or simple heuristic rule models (Gigerenzer, Todd, & the ABC Research Group, 1999; Payne, Bettman, & Johnson, 1993). It is difficult for "the new kids" to break into this field for a very important reason: They just seem too complex in comparison. Computational models are constructed from a large number of elementary units that are tightly interconnected to form a complex dynamical system. So the question, "what does this extra complexity buy us?," is raised. Computational theorists first have to prove that their models are worth the extra complexity. This chapter provides some answers to that challenge.

First, the current state of decision research applied to preferences under uncertainty is reviewed. The evolution of the algebraic utility approach that has dominated the field of decision making is described, showing a steady progression away from a simple and intuitive principle of maximizing expected value. The development of utility theories into their current form has included modifications for the subjective assessment of objective value and probability, with the most recent work focusing on finer specification of the latter. The impetus for these modifications is then discussed; in particular, specific and pervasive "paradoxes" of human choice behavior are briefly reviewed. This section arrives at the conclusion that no single utility theory provides an accurate descriptive model of human choice behavior.

Then, computational approaches to decision making are introduced, which seem more promising in their ability to capture robust trends in human choice behavior. This advantage is due to their common focus on the micro-mechanisms of the underlying deliberation process, rather than solely on the overt choice behavior driven by choice stimuli. A number of different approaches are introduced, providing a broad survey of the current corpus of computational models of decision making. The fourth section focuses on one particular model to offer a detailed example of the computational approach. Specifically, decision field theory is discussed, which has benefit from the most extensive (to date) application to a variety of choice domains and empirical phenomena.

The fifth section provides concrete illustration of how the computational approach can account for all of the behavioral paradoxes in the second section that have contested utility theories. Again, decision field theory is recruited for this analysis because of its success in accounting for all the relevant phenomena. However, the extent to which the other computational models have been successful in accounting for the results is also discussed. We conclude with comparisons among the computational models introduced, summary comparisons between the computational approach, and utility-based models of decision making.

2. Decision Models: State of the Art

2.1. *The Evolution of Utility-Based Models*

Decision theory has a long history, starting as early as the seventeenth century with probabilistic theories of gambling by Blaise Pascal and Pierre Fermat. Consider an option, or prospect, that offers some n number of quantifiable outcomes, $\{x_1, \ldots, x_n\}$, each with some specified probability, $\{p_1, \ldots, p_n\}$, respectively. The initial idea was that the decision maker should choose to maximize the long run average value or *expected value* (EV), $EV = \sum p_j \cdot x_j$. But the EV principle soon came under attack because it prescribes paying absurd prices to play a celebrated gamble known as the St. Petersburg paradox. It was also criticized because it fails to explain why people buy insurance (the premium exceeds the expected value). To fix these problems, Daniel Bernoulli (1738) proposed that the objective outcome x_j be replaced with the subjective utility of this outcome $u(x_j)$, and recommended that the decision maker should choose to maximize the *expected utility* (EU), $EU = \sum p_j \cdot ux_j)$.

For many years, Bernoulli's EU theory was disregarded by economists because it lacked a rational or axiomatic foundation. For example, why should one choose on the basis of expectation if the game is played only once? Von Neumann and Morgenstern (1947) rectified this problem by (a) proposing a set of rational axioms (e.g., transitivity, independence, solvability), and (b) proving that the EU principle uniquely satisfies these axioms. This led to EU theory being accepted by economists as the rational basis for making decisions. Thus far, EU theory was restricted to decisions with objectively known probabilities (e.g., well-defined lotteries). Shortly afterward, Savage (1954) provided an axiomatic foundation for assigning personal probabilities to uncertain events (e.g., presidential elections).

Unfortunately, people are not always rational, and subsequent empirical research soon demonstrated systematic violations of these rational axioms (see Allais, 1961; Ellsberg, 1953). To explain these violations, Kahneman and Tversky (1979) developed *prospect theory*, which changed EU theory in two important ways. Following an earlier suggestion by Edwards (1962), they replaced the objective probabilities p_i with subjective decision weights $\pi(p_i)$, where π is an inverse S shaped function. Unlike Savage's (1954) theory, these decision weights are not constrained to obey the laws of probability. Second, the utility function was defined with respect to a reference point: for losses (below the reference), the function is convex (risk seeking); for gains (above the reference), the function is concave (risk averse); and the function is steeper on the loss compared with the gain side (loss aversion). The initial prospect theory was severely criticized for two main reasons (see Starmer, 2000): (1) it predicted preferences for stochastically dominated options that are never empirically observed (anomalies that had to be removed by ad hoc editing operations); and (2) the theory was limited to binary outcomes, and it broke down and made poor predictions for a larger number of outcomes (Lopes & Oden, 1999).

Recognizing these limitations, Tversky and Kahneman (1992) modified and extended prospect theory to form cumulative prospect theory (CPT), which builds on earlier ideas of rank dependent utility (RDU) theories (Quiggin, 1982). The problem to be solved was the following: On the one hand, nonlinear decision weights were needed to explain violations of the rational axioms; but on the other hand, nonlinear transformations of outcome probabilities led to absurd predictions. To overcome this problem, RDU theories such as CPT employ a more sophisticated method for computing decision weights.[1] Suppose payoffs are rank-ordered in preference according to the index j so $u(x_{j+1}) > u(x_j)$. The rank dependent decision weight for outcome x_j is then defined by the formula: $w(x_j) = \pi(\sum_j^n p_j) - \pi(\sum_{j+1}^n p_j)$ for $j = n-1,\ n-2, \ldots, 2, 1$, and $w(x_n) = \pi(p_n)$.

Here, π is a monotonically increasing weight function designed to capture optimistic (more weight to higher outcomes) or pessimistic (more weight to lower outcomes) beliefs of a decision maker. The term $(\sum_j^n p_j)$ is called the decumulative probability (one minus the cumulative probability), which is the probability of getting a payoff at least as good as x_j. Whereas prospect theory transformed the outcome probabilities, $\pi(p_j)$, CPT transforms the decumulative probabilities, $\pi(\sum_j^n p_j)$. By doing this, one can account for systematic violations of the EU axioms, while at the same time avoid making absurd predictions about dominated options. This is the current state of utility theories.

2.2. *Problems with Utility Models: Paradoxes in Decision Making*

This section briefly and selectively reviews some important paradoxes of decision making (for a more complete review, see Rieskamp, Busemeyer, & Mellers, 2006; Starmer, 2000) and points out shortcomings of utility theories in explaining these phenomena.

1 Note that CPT is one exemplar from the class of RDU, which in turn are a subset of the more general EU approach. For the current chapter, reference to one class subsumes the more specific model(s); e.g., claims regarding RDU theory apply also to CPT.

2.2.1. ALLAIS PARADOX

This most famous paradox of decision making (Allais, 1979; see also Kahneman & Tversky, 1979) was designed to test expected utility theory. In one example, the following choice was given:

A: "win \$1 M (million) dollars for sure,"
B: "win \$5 M with probability .10, or \$1 M with probability .89, or nothing."

Most people preferred prospect A even though prospect B has a higher expected value. This preference alone is no violation of expected utility theory – it simply reflects a risk averse utility function. The violation occurs when this first preference is compared with a second preference obtained from a choice between two other prospects:

A′: "win \$1 million dollars with probability .11, or nothing,"
B′: "win \$5 million dollars with probability .10, or nothing."

Most people preferred prospect B', and the (A, B') preference pattern is the paradox.

To see the paradox, one needs to analyze this problem according to expected utility theory. These prospects involve a total of three possible final outcomes: $\{x_1 = \$0,\ x_2 = \$1\text{ M},\ x_3 = \$5\text{ M}\}$. Each prospect is a probability distribution, $(p_1,\ p_2,\ p_3)$, over these three outcomes, where p_j is the probability of getting payoff x_j. Thus, the prospects are:

$$A = (0, 1, 0) \qquad A' = (.89, .11, 0)$$
$$B = (.01, .89, .10) \qquad B' = (.90, 0, .10).$$

Now define three new prospects:

$$O = (0, 1, 0) \qquad Z = (1, 0, 0)$$
$$F = (1/11, 0, 10/11).$$

It can be seen that $A = (.11) \cdot O + (.89) \cdot O$ and $B = (.11) \cdot F + (.89) \cdot O$, producing $EU(A) - EU(B) = [(.11) \cdot EU(O) + (.89) \cdot EU(O)] - [(.11) \cdot EU(F) + (.89) \cdot EU(O)]$.

The common branch, $(.89) \cdot EU(O)$, cancels out, making the comparison of utilities between A and B reduce to a comparison of utilities for O and F. It can also be seen that:

$A' = (.11) \cdot O + (.89) \cdot Z$ and $B' = (.11) \cdot F + (.89) \cdot Z$, producing $EU(A') - EU(B') = [(.11) \cdot EU(O) + (.89) \cdot EU(Z)] - [(.11) \cdot EU(F) + (.89) \cdot EU(Z)]$.

Again a common branch, $(.89) \cdot EU(Z)$, cancels out, making the comparison between A' and B' reduce to the same comparison between O and F. More generally, EU theory requires the following *independence axiom*: for any three prospects $\{A, B, C\}$, if A is preferred to B, then $A' = p \cdot A + (1 - p) \cdot C$ is preferred to $p \cdot B + (1 - p) \cdot C = B'$. The Allais preference pattern (A, B') violates this axiom.

To account for these empirical violations, the independence axiom has been replaced by weaker axioms (see Luce, 2000, for a review). The new axioms have led to the development of the RDU class of theories introduced earlier, including CPT, which can account for the Allais paradox. However, the RDU theories (including CPT) must satisfy another property called stochastic dominance.

2.2.2. STOCHASTIC DOMINANCE

Assume again that the payoffs are rank ordered in preference according to the index j, so $u(x_{j+1}) > u(x_j)$. Define X as the random outcome produced by choosing a prospect. Prospect A stochastically dominates prospect B if and only if $\Pr[u(X) \geq u(x_j) \mid A] \geq \Pr[u(X) \geq u(x_j) \mid B]$ for all x_j.

In other words, if A offers at least as good a chance as B of obtaining each possible outcome or better, then A stochastically dominates B.[2] The reason RDU theories (e.g., CPT) must satisfy stochastic dominance (predict choice of stochastically dominating prospects) is straightforward. If A stochastically dominates B with respect to the payoff probabilities, then it follows that A stochastically dominates B with respect to the decision weights, which implies that the RDU for A is greater than that for B, and this finally implies that A is preferred to B. Unfortunately for decision theorists, human preferences do not obey this property either – systematic violations of stochastic dominance have been reported (Birnbaum & Navarrete, 1998; Birnbaum, 2004). In one example, the following choice was presented:

F: "win \$98 with .85, or \$90 with .05, or \$12 with .10,"
G: "win \$98 with .90, or \$14 with .05, or \$12 with .05."

Most people chose F in this case, but it is stochastically dominated by G. To see this, we can rewrite the prospects as follows:

F′: "win \$98 with .85, or \$90 with .05, or \$12 with .05, or \$12 with .05,"
G′: "win \$98 with .85, or \$98 with .05, or \$14 with .05, or \$12 with .05."

Most people chose G' in this case. The choice of F violates the principle of stochastic dominance, which is contrary to RDU theories such as CPT. More complex decision weight models, such as Birnbaum's Tax model, are required to not only explain violations of stochastic dominance, but to simultaneously account for the pattern (F, G′; see Birnbaum, 2004).

2.2.3. PREFERENCE REVERSALS

Violations of independence and stochastic dominance are two of the classic paradoxes of decision making. Perhaps the most serious challenge for all utility theories is one that calls into question the fundamental concept of preference. According to most utility theories (including prospect theory), there are two equally valid methods for measuring preference – one based on choice, and a second based on price. If prospect A is chosen over prospect B, then $u(A) > u(B)$, which implies that the price equivalent for prospect A should be greater than the price equivalent for prospect B (this follows from the relations, $\$_A = A > B = \$_B$, where $\$_K$ is the price equivalent of prospect K). Contrary to this fundamental prediction, systematic reversals of preferences have been found between choices

2 Note that, technically, A must also offer a better chance of obtaining at least one outcome. That is, the inequality must be strict for at least one outcome, otherwise the prospects A and B are identical.

and prices (Grether & Plott, 1979; Lichtenstein & Slovic, 1971; Lindman, 1971; Slovic & Lichtenstein, 1983). In one example, the following prospects were presented:

P: "win $4 with 35/36 probability,"
D: "win $16 with 11/36 probability."

Most people chose prospect *P* over prospect *D*, even though *D* has a higher expected value – they tend to be risk averse with choices. The same people, however, most frequently gave a higher price equivalent to prospect *D* than to prospect *P*. Furthermore, another interesting finding in need of explanation is that the variance of the prices for prospect *D* is much larger than that for prospect *P* (Bostic, Herrnstein, & Luce, 1990).

Tversky, Sattath, & Slovic (1988) initially explained preference reversals between choice and price by arguing that decision makers place more weight on the probability dimension when making choices, whereas the price task shifts weight to the price dimension. Alternatively, Mellers, Schwartz, and Cooke (1998) argued that decision makers use different strategies when making choices versus prices. However, a serious problem for both of these explanations is that preferences also reverse when individuals are asked to give two different types of prices, such as minimum selling prices (willingness to accept [WTA]) versus maximum buying prices (willingness to pay [WTP]), for the same prospects (Birnbaum & Zimmerman, 1998). Consider the following two prospects:

F: "win $60 with probability .50, otherwise $48."
G: "win $96 with probability .50, otherwise $12."

People gave a higher WTA for prospect *G* compared with prospect *F*, but the opposite order was found for WTP. So, not only do preferences change depending on whether choices or prices are used, but also when different types of prices are used. Furthermore, such violations extend beyond trivial tasks involving hypothetical or low-stakes gambles to situations involving more realistic consequences, such as managerial decisions, medical decisions, environmental protection policies, and highway safety programs.

Neither choice-pricing nor WTP-WTA reversals can be explained with a single utility model such as prospect theory, but only by assuming arbitrary task-dependent changes in the decision weights and/or utility function and/or combination of weight and utility. These unnerving findings have led researchers to question stability of preferences and to argue instead that preferences are constructed on the fly in a task-dependent manner (e.g., Slovic, 1995).

2.2.4. CONTEXT-DEPENDENT PREFERENCES

A final challenge for utility theories is that preferences seem to depend not only on changes in the task, but also in changes in the context produced by the choice set for a single task. These preference reversals involve violations of a principle called *independence from irrelevant alternatives*. According to this principle, if option *A* is chosen most frequently over option *B* in a choice set that includes only $\{A, B\}$, then *A* should be chosen more frequently over *B* in a larger choice set $\{A, B, C\}$ that includes a new option C. This principle is required by a large class of utility models called simple scalable utility models (see Tversky, 1972). However, empirical evidence points to at least three direct violations of this principle.

The first violation is produced by what is called the similarity effect (Tversky, 1972; Tversky & Sattath, 1979), in which case the new option, labeled *S*, is designed to be similar and competitive with the common option *B*. In one example, participants chose among hypothetical candidates for graduate school that varied in terms of intelligence and motivation scores:

Candidate A: Intelligence = 60, Motivation = 90,
Candidate B: Intelligence = 78, Motivation = 25,

Candidate S: Intelligence = 75, Motivation = 35.

Participants chose *B* more frequently than *A* in a binary choice. However, when candidate *S* was added to the set, then preferences reversed and candidate *A* became the most popular choice. The similarity effect rules out all simple scalable utility models, but it can be explained by a heuristic choice model called the elimination by aspects (EBA) model (Tversky, 1972). According to this model, decision makers sample a feature based on its importance and eliminate any option that does not contain the selected feature; the process continues until there is only one option left, and the last surviving option is then chosen. Applying EBA to the previous example, if grade-point average is most important, then *A* is most likely to be eliminated at the first stage, leaving *B* as the most frequent choice; however, when *S* is added to the set, then both *B* and *S* survive the first elimination, and *S* reduces the share of *B*.

The second violation is produced by what is called the attraction effect (Huber, Payne, & Puto, 1982; Huber & Puto, 1983; Simonson, 1989), in which case the new option, labeled *D*, is similar to *A* but dominated by *A*. In one example, participants chose among cars varying in miles per gallon and ride quality:

Brand A: 73 rating on ride quality, 33 miles per gallon (mpg),
Brand B: 83 rating on ride quality, 24 mpg,
Brand D: 70 rating on ride quality, 33 mpg.

Brand *B* was more frequently chosen over brand *A* on a binary choice; however, adding option *D* to the choice set reversed preferences so that brand *A* became most popular. In this second case, the new option helps rather than hurts the similar option. The attraction effect is important because it violates another principle called *regularity*, which states that adding an option to the set can never increase the popularity of one of the original options from the subset. The EBA model satisfies regularity, and therefore it cannot explain the attraction effect (Tversky, 1972).

The third violation is produced by what is called the compromise effect (Simonson, 1989; Simonson & Tversky, 1992), in which a new extreme option *A* is added to the choice set. In one example, participants chose among batteries varying in expected life and corrosion rate:

Brand A: 6% corrosion rate, 16 hours duration,
Brand B: 2% corrosion rate, 12 hours duration,
Brand C: 4% corrosion rate, 14 hours duration.

When given a binary choice between *B* and C, brand *B* was more frequently chosen over brand C. However, when option *A* was added to the choice set, then brand C was chosen more often than brand *B*. Thus, adding an extreme option *A*, which turns option C into a compromise, reverses the preference orders obtained between the binary and triadic choice methods. The compromise effect is interesting because it rules out another heuristic choice rule called the lexicographic (LEX), or "take the best," strategy. According to this strategy, the decision maker first considers the most important dimension and picks the best alternative on this dimension, but if there is a tie, then decision maker turns to the second most important dimension and picks the best on this dimension, and so forth. According to the LEX strategy, individuals should never choose the compromise option!

The collection of results presented in this section indicate that preferences among a set of options are not subject to the calculus of probability and are dependent on the choice context and the elicitation method. These results are only a subset of the decades of research showing that human decisions do not correspond to those predicted by utility models. Any serious model of decision making must account for effects such as the robust and representative examples mentioned in this section. We now turn to examining a distinctly different type of modeling approach that shows promise in this respect.

3. Computational Models of Decision Making: A Survey

In an attempt to retain the basic utility framework, constraints on utility theories are being relaxed, and the formulas are becoming more deformed. Recently, many researchers have responded to the growing corpus of phenomena that challenge traditional utility models by applying wholly different approaches. That is, rather than continuing to modify utility equations to accommodate each new empirical trend, these researchers have adopted alternative representations of human decision making. The common thread among these approaches is their attention to the processes, or computations, that are assumed to produce observable decision behavior. Beyond this, the popular approaches outlined in this section diverge in precisely how they model decision making.

3.1. *Heuristic Rule-Based Systems*

Payne, Bettman, and Johnson (1992, 1993) propose an adaptive approach to decision making. Essentially, this approach assumes that decision makers possess a repertoire of distinct decision strategies that they may apply to any given task. The repertoire of strategies usually includes noncompensatory rules that do not require trade-offs among attributes, such as EBA and LEX, as well as compensatory rules that are based on attribute trade-offs such as a weighted additive (WADD) rule or EU rule. Furthermore, it is assumed that the strategy applied is selected as a trade-off between the mental effort required to apply the strategy and the accuracy or performance of the strategy. Thus, in trivial situations or those involving extreme time pressure, individuals may employ relatively simple strategies that do not involve complex calculations such as the LEX or EBA rules. In contrast, in important situations where a high level of performance is required, decision makers may apply more cognitively intensive strategies such as the WADD or EU rule.

This approach assumes that each possible strategy is assembled from elementary information processing units, such as “retrieve,” “store,” “move,” “compare,” “add,” “multiply,” and so forth. (Payne et al., 1993). For example, the EBA rule might be instantiated by a “retrieve” of a prospect’s attribute value, followed by a “compare” to some threshold value defining deficiency. EU could be formalized by a “multiply” of subjective probability and utility values, the “store” of each product, and an “add” across products; choice is defined by a “compare” operation among expected utilities. Mental effort is defined by the sum of processing times for these elementary mental operations, and accuracy is typically defined by performance relative to the WADD or EU rule.

Gigerenzer and colleagues (Gigerenzer et al., 1999) have developed a closely related approach. Their simple heuristics are formulated in terms of their rules for (a) searching through information, (b) stopping this search, and (c) selecting an option once the search concludes. For example, Brandstätter, Gigerenzer, and Hertwig (2006) recently proposed a LEX model called the “priority heuristic,” which assumes the following process for positively valued gambles: (1) first compare the lowest outcomes for each prospect, and if this difference exceeds a cutoff, then choose the best on this comparison; otherwise (2) compare the probabilities associated with the lowest payoffs, and if this difference exceeds a cutoff then choose the best on this comparison; otherwise (3) compare the maximum possible payoff for each prospect and choose the best on this maximum.

The strength of heuristic models is their ability to explain effects of effort, conflict, time pressure, and emotional content on choices and other processing measures (e.g., amount of information searched, order of search) in terms of changes in decision strategies. However, one drawback to these models is their lack of specification across applications; it is often difficult to determine exactly which strategy is used in any given situation. Furthermore, when considering the findings summarized earlier, the heuristic models cannot account for the all of these results reviewed previously despite this flexibility. They have been used to explain violations of independence for risky choices but not the violations of stochastic dominance. They also have been used to explain preference reversals between choice and prices, but not between buying and selling prices. Finally they can explain the similarity effect but not the compromise or attraction effect.

3.2. *Dynamic Systems/Connectionist Networks*

Many researchers prefer to adopt a single dynamic process model of decision making rather than proposing a tool box of strategies. This idea has led to the development of several computational models that are formulated as connectionist models or dynamic systems (see Chapter 2 on connectionist models and Chapter 4 on dynamic systems in this volume).

3.2.1. AFFECTIVE BALANCE THEORY

Grossberg and Gutowski (1987) presented a dynamic theory of affective evaluation based on an opponent processing network called a gated dipole neural circuit. Habituating transmitters within the circuit determine an affective adaptation level, or reference point, against which later events are evaluated. Neutral events can become affectively charged either through direct activation or antagonistic rebound within the habituated dipole circuit. This neural circuit was used to provide an explanation for the probability weighting and value functions of Kahneman and Tversky's (1979) prospect theory, and preference reversals between choices and prices. However, this theory cannot explain preference reversals between buying and selling prices, nor can it explain violations of stochastic dominance. Finally, the affective balance theory has never been applied to more than two choice options, so it is not clear how it would explain the similarity, attraction, and compromise context effects.

3.2.2. ECHO

Holyoak and Simon (1999) and Guo and Holyoak (2002) proposed a connectionist network, called ECHO, adapted from Thagard and Millgram (1995). According to this theory, there is a special node, called the external driver, representing the goal to make a decision, which is turned on when a decision is presented. The driver node is directly connected to attribute nodes, with a constant connection weight. Each attribute node is connected to an alternative node with a bidirectional link, which allows activation to pass back and forth from the attribute node to the alternative node. The connection weight between an attribute node and an alternative node is determined by the value of the alternative on that attribute. There are also constant lateral inhibitory connections between the alternative nodes to produce a competitive recurrent network.

The decision process works as follows. On presentation of a decision problem, the driver is turned on and applies constant input activation into the attribute nodes, and each attribute node then activates each alternative node (differentially depending on value). Then each alternative node provides positive feedback to each attribute node and negative feedback to the other alternative nodes. Activation in the network evolves over time according to a nonlinear dynamic system, which keeps the activations bounded between zero and one. The decision process stops as soon as the changes in activations fall below some threshold. At that point, the probability of choosing an option is determined by a ratio of activation strengths.

The ECHO model has been shown to account for the similarity and attraction effect, but it cannot account for the compromise effect. It has not been applied to risky choices, so it remains unclear how it would explain violations of independence or stochastic dominance. Finally, this theory is restricted to choice behavior, and it has no mechanisms for making predictions about prices. One interesting prediction of the ECHO model is that the weight of an attribute changes during deliberation in the direction of the currently favored alternative. Evidence supporting this prediction was reported by Simon, Krawczyk, and Holyoak (2004).

3.2.3. LEAKY COMPETING ACCUMULATOR MODEL

Usher and McClelland (2004) proposed a connectionist network model of decision making called the leaky competing accumulator model. Preference is based on the sequential evaluation of attributes, where each evaluation compares the relative advantages and disadvantages of each prospect. These comparisons are integrated over time for each option by a recursive network. The accumulation continues until a threshold is crossed, and the first option to reach the threshold is chosen.

This theory is closely related to decision field theory (described later), with the following important exceptions. First, the activation for each option is restricted to remain positive at all times, which requires the temporal integration to be nonlinear. Second, the leaky competing accumulator model adopts Tversky and Kahneman's (1991) loss aversion hypothesis so that disadvantages have a larger impact than advantages.

Usher and McClelland (2004) have shown that the leaky competing accumulator can explain the similarity, attraction, and compromise effects using a common set of parameters. However, this model has not been applied to risky choices or to preference reversals.

3.3. *Models Cast in Cognitive Architectures*

Some researchers have taken advantage of the extensive work that has been done in developing comprehensive cognitive architectures that can then be specified for almost any conceivable individual task (see Chapter 6 on cognitive architectures in this volume). In particular, researchers have recently formulated models within two popular cognitive architectures for choice tasks that are the focus of the current chapter.

3.3.1. SUBSYMBOLIC AND SYMBOLIC COMPUTATION IN ACT-R

Although one of the most popular cognitive architectures, ACT-R, incorporates a simple expected utility mechanism by default other researchers have realized the drawbacks with the expected utility approach and developed alternative models within ACT-R. Specifically, Belavkin (2006) has developed two models that can correctly predict the Allais paradox (it has not been applied to the other paradoxes). In fact, these decision models are not unique to the ACT-R implementation proposed by Belavkin (2006); each model is actually a probabilistic extension of earlier simple heuristic rules guiding choice.

The first model essentially reduces to a simple rule of maximizing the probability of the largest outcome possible. Due to the negative correlation that typically exists between outcome and probability (e.g., to maintain constant expected value across gambles), this first rule results in the likelihood of choosing the option with the larger outcome to be equal to the probability of this outcome. The second model is formulated at the symbolic rule level in ACT-R and defines preference relations on each component of the stimuli (i.e., first outcome, probability of first outcome, second outcome, and probability of second outcome). A simple tally rule is assumed, and the proportion of total relations (including indifference) that favor each option produces the probability of choosing the option. Although each of these simple rule models can predict choices that produce the Allais paradox, they cannot predict a number of more basic results. For example, in both models, changing the value of an outcome does not affect choice if the rank order is preserved, contrary to empirical evidence.

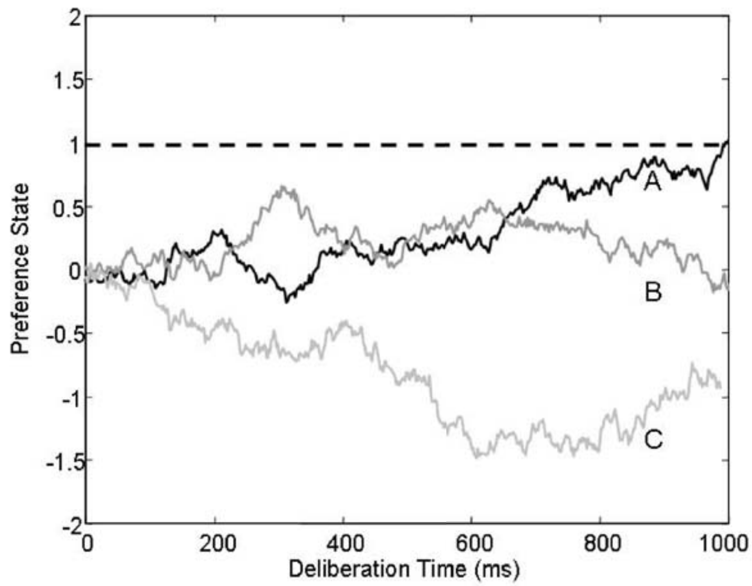

Figure 10.1. Illustration of preference evolution for three options (A, B, and C), according to decision field theory. The threshold is shown as a dashed line; the three options are shown as solid lines of different darkness.

4. Computational Models of Decision Making: A Detailed Example

It is impossible to describe all of the previously mentioned computational models in detail, so this section will focus on one, called decision field theory (DFT; Busemeyer & Townsend, 1993; Diederich, 1997; Roe, Busemeyer, Townsend, 2001; Johnson & Busemeyer, 2005a).[3] This model has been more broadly applied to decision-making phenomena compared with the other computational models at this point.

4.1. *Sequential Sampling Deliberation Process*

DFT is a member of a general class of sequential sampling models that are commonly used in a variety of fields in cognition (Ashby, 2000; Laming, 1968; Link & Heath, 1975; Nosofsky & Palmeri, 1997; Ratcliff, 1978; Smith, 1995; Usher & McClelland, 2001). The basic ideas underlying the decision process for sequential sampling models are illustrated in Figure 10.1. Suppose the decision maker is initially presented with a choice between three risky prospects, *A*, *B*, C, at time $t = 0$. The horizontal axis on the figure represents deliberation time (in milliseconds), and the vertical axis represents preference strength. Each trajectory in the figure represents the preference state for one of the risky prospects at each moment in time.

Intuitively, at each moment in time, the decision maker thinks about various payoffs of each prospect, which produces an affective reaction, or *valence*, to each prospect. These valences are integrated across time to produce the preference state at each moment. In this example, during the early stages of processing (between 200 and 300 ms), attention is focused on advantages favoring prospect *B*, but later (after 600 ms), attention is shifted toward advantages favoring prospect *A*. The stopping rule

3 The name "decision field theory" reflects the influence of Kurt Lewin's (1936) field theory of conflict.

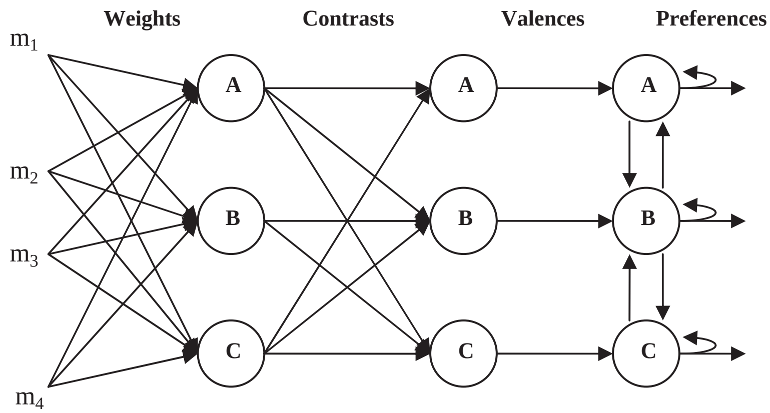

Figure 10.2. Connectionist network representation of decision field theory.

for this process is controlled by a threshold (which is set equal to 1.0 in this example): The first prospect to reach the top threshold is accepted, which in this case is prospect A after about 1 second. Choice probability is determined by the first option to win the race and cross the upper threshold, and decision time is equal to the deliberation time required by one of the prospects to reach this threshold.

The threshold is an important parameter for controlling speed-accuracy trade-offs. If the threshold is set to a lower value (about .50) in Figure 10.1, then prospect B would be chosen instead of prospect A (and done so earlier). Thus, decisions can reverse under time pressure (see Diederich, 2003). High thresholds require a strong preference state to be reached, which allows more information about the prospects to be sampled, prolonging the deliberation process and increasing accuracy. Low thresholds allow a weak preference state to determine the decision, which cuts off sampling information about the prospects, shortening the deliberation process and decreasing accuracy. There are many examples of task and individual variables that could determine the threshold for an individual application. As an example of the former, under high time pressure, decision makers must choose a low threshold; but under low time pressure, a higher threshold can be used to increase accuracy. Concerning personal variables, very careful and deliberative decision makers tend to use a high threshold, and impulsive or careless decision makers can be described as using a low threshold.

4.2. *Connectionist Network Interpretation*

Figure 10.2 provides a connectionist interpretation of DFT for the example shown in Figure 10.1. Assume once again that the decision maker has a choice among three risky prospects, and also suppose for simplicity that there are only four possible final outcomes. Thus, each prospect is defined by a probability distribution across these same four payoffs. The subjective, affective values produced by each payoff are represented by the inputs, m_j, shown on the far left side of this network. At any moment in time, the decision maker anticipates the payoff of each prospect, which produces a momentary evaluation, $U_i(t)$, for prospect i, shown as the first layer of nodes in Figure 10.2. This momentary evaluation is an attention-weighted average of the affective evaluation of each payoff: $U_i(t) = \sum W_{ij}(t) \cdot m_j$. The attention weight at time t, $W_{ij}(t)$, for payoff j offered by prospect i, is assumed to fluctuate according to a stationary stochastic process. This reflects the idea that attention is shifting from moment to moment, causing changes in the anticipated payoff of each prospect across time.

The momentary evaluation of each prospect is compared with other prospects to form a valence for each prospect at each moment, $v_i(t) = U_i(t) - U_.(t)$, where $U_.(t)$ equals the average momentary evaluation across all the prospects. The valence $v_i(t)$ represents the relative advantage or disadvantage of prospect i at time t, and this is shown as the second layer of nodes in Figure 10.2. The total valence balances out to zero so that all the options cannot become attractive simultaneously.

Finally, the valences are the inputs to a dynamic system that integrates the valences over time to generate the output preference states. The output preference state for prospect i at time t is symbolized as $P_i(t)$, which is represented by the last layer of nodes in Figure 10.2 (and plotted as the trajectories in Figure 10.1). The dynamic system is described by the following linear stochastic difference equation for a small time step h in the deliberation process:

$$P_i(t+h) = \sum\nolimits_j s_{ij} \cdot P_j(t) + v_i(t+h) \tag{10.1}$$

The positive self-feedback coefficient, $s_{ii} = s > 0$, controls the memory for past input valences for a preference state. Values of $s_{ii} < 1$ suggest decay in the memory or impact of previous valences over time, whereas values of $s_{ii} > 1$ suggest growth in impact over time (primacy effects). The negative lateral feedback coefficients, $s_{ij} = s_{ji} < 0$ for $i \neq j$, produce competition among actions so that the strong inhibit the weak. In other words, as preference for one prospect grows stronger, then this moderates the preference for other prospects. The magnitudes of the lateral inhibitory coefficients are assumed to be an increasing function of the similarity between choice options. These lateral inhibitory coefficients are important for explaining context effects on preference.

Formally, this decision process is a Markov process, and matrix formulas have been mathematically derived for computing the choice probabilities and distribution of choice response times (for details, see Busemeyer & Diederich, 2002; Busemeyer & Townsend, 1992; Diederich & Busemeyer, 2003). Alternatively, Monte Carlo computer simulation can be used to generate predictions from the model.

4.3. *Attention Switching Mechanism*

What is the psychological source of decision weights? According to DFT, an attention process is used to generate the predicted payoff for each prospect at each time step of the sequential sampling process. In this context, the decision weight for a payoff equals the average amount of time an individual spends paying attention to that payoff. Consequently, the decision weights are derived from a micro-process model of attention (Johnson & Busemeyer, 2006).

Consider a prospect with payoffs $x_1 \leq x_2, \ldots, \leq x_n$ and associated probabilities $(p_1, \ldots, p_n)$. The attention process starts at the lowest payoff and works its way up the ranks. Given that the attention process is focused on a particular payoff x_j for $1 < j < n$, it can make four transitions: predict x_j with probability p_j; do not predict this right away, but remain focused on it with probability $\beta \cdot (1 - p_j)$; or switch the focus up to the next highest payoff or down to the next lowest payoff with equal probability, $(1 - \beta) \cdot (1 - p_j)/2$. If attention is focused on the lowest (highest) payoff, then focus may only switch to the next lowest (highest) payoff; that is, the probability of switching focus is $(1 - \beta) \cdot (1 - p_j)$, for $j = \{1, n\}$. This attention mechanism is then used to mathematically derive (again using Markov chain theory) the mean attention weights, $w_{ij} = \mathrm{E}[W_{ij}(t)]$, for DFT (see Johnson & Busemeyer, 2006). In this way, all of the decision weight parameters are derived on the basis of a single attention parameter, $0 \leq \beta \leq 1$, that represents the tendency to dwell on any given outcome once focused on the outcome.

4.4. *Response Mechanism*

How can a choice process be used to determine prices, yet still produce preference reversals? According to DFT, a sequential comparison process is used to search and find a price that makes the decision maker indifferent when faced with a choice between a prospect and a price (Johnson & Busemeyer, 2005a).

Consider, for example, the task of finding a price for the *D* bet given earlier, "win \$16 with probability 11/36." For simplicity, assume the feasible set of candidate prices includes the dollar values \$0, \$1, \$2, . . . , \$16. For a simple price equivalent, the most efficient place to start searching is in the middle of this set (\$8); when buying, it is advantageous to start bargaining with the lowest possible bid (\$0); and it is advantageous for sellers to start by asking for the highest price (\$16). The sequential comparison then inserts this starting value into a binary choice process (the *D* prospect is compared with the candidate dollar value). This comparison process can result in one of three outputs: (a) if the process results in (implicit) choice favoring the prospect *D* over the candidate value, then the price is too low, and it is incremented by a dollar; (b) if the process results in preference for the candidate value over the prospect *D*, then the price is too high, and the price is reduced by a dollar; however, (c) each time that the comparison process transits through the zero (indifference) preference state, then there is some probability, r, that the comparison process will stop and exit, and report finding a price equivalent. This sequential comparison process is then used to mathematically derive (again using Markov chain theory) the entire distribution of prices for gambles (see Johnson & Busemeyer, 2005a).

4.5. *Model Parameters*

It is now possible to identify and compare the parameters of DFT model with those of RDU theories, such as CPT. First, DFT has a set of affective values, m_j, that correspond to the utilities of outcomes, $u(x_j)$, used in RDU theories. Second, DFT has a set of mean attention weights, w_{ij}, that correspond to the decision weights, $w_i(x_j)$, of RDU theories. However, the weights for DFT are generated from an attention mechanism, which requires only one parameter, β. To account for prices, DFT requires only one additional parameter, the exit rate parameter r, whereas RDU theories require a new set of weights for choices and prices to account for preference reversals.

In addition, DFT includes three types of parameters to describe properties of human decision making that RDU models (including CPT) cannot. First, DFT uses a threshold-bound parameter to account for speed-accuracy trade-offs (RDU theories fail to do this because they are static). Second, DFT includes a variance term to account for the probabilistic nature of choice (RDU theories are deterministic, and probabilistic extensions require additional parameters). A parameter for the self-feedback coefficient, $s_{ii} = s$, is needed to account for primacy/recency effects on the growth of preferences over time, and parameters for the lateral inhibition coefficients, $s_{ij} = s_{ji}$ for $i \neq j$, are needed to explain context-dependent preferences.

5. Accounting for Paradoxes in Decision Making

As indicated by the selective survey of results in Section 2.2, human decision-making behavior is complex, even under extremely simple decision situations. Can the computational models account for this daunting collection of empirical results? In this section, we will show how DFT is able to account for all of the findings introduced in Section 2.2. Although this is the only theory that has been shown to account for this entire collection of results, we also mention where appropriate the success or failure of other computational approaches in accounting for some of these findings.

5.1. *Accounting for Violations of Independence and Stochastic Dominance*

Recall that RDU theories (including CPT) are unable to account for violations of stochastic dominance. The attention-switching mechanism of DFT is responsible for its ability to predict violations of

Table 10.1: Predictions derived from micro-process model of attention to payoffs

Prospect	*Probabilities*	*Weights*	*Mean value*
Allais problem			
A	0, 1, 0	0, 1, 0	1.00
B	.01, .89, .10	.03, .96, .01	.986
A′	.89, .11, 0	.99, .01, 0	.011
B′	.90, 0, .10	.99, 0, .01	.045
Stochastic dominance problem			
F	.10, .05, .85	.40, .16, .44	62.65
G	.05, .05, .90	.24, .20, .56	60.64
F′	.05, .05, .05, .85	.27, .28, .12, .33	49.85
G′	.05, .05, .05, .85	.27, .28, .12, .33	51.38

independence and stochastic dominance (see Johnson & Busemeyer, 2006). Table 10.1 presents the predictions for both the Allais and the stochastic dominance choice problems from Section 2.2, when the "dwell parameter" was set to $\beta = .70$. The columns show the prospect, the probabilities, the weights, and the mean values (assuming $\mathrm{E}[U_i(t)] = u_i = \sum w_{ij} \cdot m_j$ with $m_j = x_j$). As can be seen in this table, both paradoxes are explained using the same attention mechanism and the same parameter value. Intuitively, the tendency to begin by considering low outcomes, coupled with a moderate dwelling probability, results in "overweighting" of the small probabilities associated with the lowest outcomes of the prospects. Note that the β parameter and/or m_j values could be fit to amplify or moderate the effects shown in Table 10.1 However, we avoid this in order to illustrate that a simple and consistent application can produce the paradox. Furthermore, Johnson and Busemeyer (2006) show how the attention process accounts for several other findings that are not reviewed here, using the same assumptions and parameter value.

5.2. *Accounting for Preference Reversals*

As noted earlier, strategy switching between tasks can explain reversals between choices and prices, but they fail to explain reversals between buying and selling prices. To illustrate the predictions of the DFT model for reversals between choice and pricing, consider prospects P and D introduced in Section 2.2.3, for which robust preference reversals have been observed. The first result that must be predicted is the risk-averse tendency found with choices (a higher proportion of P choices). To obtain this, Johnson and Busemeyer (2005a) assumed the affective values of the payoffs to be a concave function of the payoffs (specifically, $m_j = x_j^{0.7}$). This produces a higher predicted choice probability (0.68) for prospect P compared with prospect D. To generate price equivalents, the exit rate parameter for indifference was set equal to $r = .02$. This generates both a higher predicted mean price for prospect D ($4.82) compared with prospect P ($3.42), as well as a larger predicted variance in the prices for prospect D ($4.13) compared with prospect P ($.31).

Next, consider the application to prospects F and G described in Section 2.2.3. Using exactly the same parameter values and assumptions as those applied to P and D produces the following results: the mean buying price for prospect F ($52) exceeds that for prospect G ($38), but the mean selling price for prospect G ($64) is higher than that for prospect F ($56). More generally, this sequential comparison process is able to reproduce the observed preference orders for five different measures of preference

Table 10.2: Choice probabilities predicted by decision field theory for similarity, attraction, and compromise effects

Similarity		*Attraction*		*Compromise*	
Options	*Probability*	*Options*	*Probability*	*Options*	*Probability*
A: (1.0, 3.0)	.39	A: (1.0, 3.0)	.59	A: (1.0, 3.0)	.31
B: (3.0, 1.0)	.31	B: (3.0, 1.0)	.40	B: (3.0, 1.0)	.25
S: (2.99, 1.01)	.30	D: (1.0, 2.5)	.01	C: (2.0, 2.0)	.44

Note: Simulation results based on 10,000 replications.

(see Johnson & Busemeyer, 2005a): choices, price equivalents, minimum selling prices, maximum buying prices, and probability equivalents.

5.3. *Accounting for Context Dependent Preferences*

Can a single theory account for similarity, attraction, and compromise effects, using a common set of assumptions and a single set of parameter values? Recall that simple scalable utility models fail to explain the similarity effect, the EBA model fails to account for the attraction effect, and the LEX model fails to account for the compromise effect. Roe et al. (2001) initially demonstrated that DFT provides a robust and comprehensive account for all three effects. For multi-attribute choice tasks, attention is assumed to drift back and forth between attributes across time (Diederich, 1997). For example, when choosing among consumer products, attention shifts between thinking about quality and price. Although mathematical formulas have been derived for calculating the model predictions for this process (see Diederich, 1997; Roe et al., 2001), it is simpler (albeit slower) to generate predictions from computer simulations, especially when the number of alternatives is large.[4]

Predictions from DFT for an example of all three context effects are presented in Table 10.2. The values of the alternatives on each attribute are shown in the table (these determine the inputs, m_{ij} , for the network). For all three effects, the same set of parameters were used: the mean attention weight for the two attributes was set equal to .51 and .49 (reflecting slightly greater weight on the first dimension); the threshold bound was set equal to 12; the variance parameter for the valence was set equal to 1; the self-feedback coefficient was set equal to .93; the lateral inhibitory coefficient connection between the two most extremely different options, *A* and *B*, was set to zero; and the lateral inhibitory coefficient between similar option pairs was set to −.07.

Option *B* tends to be chosen more frequently in a binary choice (.55 for *B* for all three conditions), because of the larger weight given to the first attribute. However, as Shown in Table 10.2, this preference is reversed by the introduction of a third option in the triadic choice sets. As shown in Table 10.2, the model successfully reproduces all three effects: for the similarity effect, the addition of a new similar competitive option hurt option B; for the attraction effect, the addition of a new similar dominated option helped option A; and for the compromise effect, the addition of the extreme option made the compromise option C most popular.

According to DFT, the attention switching mechanism is crucial for producing the similarity effect, but the lateral inhibitory connections are critical for explaining the compromise and attraction effects. If the attention switching process is eliminated, then the similarity effect disappears, and if the lateral connections are all set to zero,

4 The predictions in Table 10.2 were generated from a simulation program available at http://mypage.iu.edu/~jbusemey/lab/sim_mdf.m.

then the attraction and compromise effects disappear. This property of the theory entails an interesting prediction about the effects of time pressure on preferences. The contrast effects produced by lateral inhibition require time to build up, which implies that the attraction and compromise effects should become larger under prolonged deliberation (see Roe et al., 2001). Alternatively, if context effects are produced by switching from a weighted average rule under binary choice to a quick heuristic strategy for the triadic choice, then these effects should get larger under time pressure. Empirical tests show that prolonging the decision process indeed increases the effects (Simonson, 1989) and time pressure decreases the effects (Dhar, Nowlis, & Sherman, 2000).

6. Discussion

This chapter began with a challenge to computational models: What can they contribute that goes beyond the explanatory power of the more popular approaches to decision making based on algebraic utility or heuristic rules? Following, a synopsis is provided that is based on the detailed discussions presented in the earlier sections. The issue of complexity of computational models, is also addressed followed by a discussion some connections to work on computational models in other domains of judgment.

6.1. *Comparison Among Models*

Modern rank dependent utility theories, such as cumulative prospect theory, are able to explain some old paradoxes of risky decision making, such as the Allais paradox. But they fail to explain new paradoxes of risky decision making, such as stochastic dominance violations. Furthermore, they cannot explain preference reversals between choice, and prices without postulating entirely new utility functions for each measure. Finally, they are unable to account for context effects on choice including similarity, attraction, and compromise effects.

Simple heuristic rule-based models allow for changes in strategy from compensatory rules (e.g., WADD or EU) to noncompensatory rules (e.g., EBA and LEX). These switches occur under time pressure or with increases in choice set size and may depend on the response measure. Simple heuristic rules can explain the Allais paradox with risky decisions, but not violations of stochastic dominance. Strategy switching between response measures can account for preference reversals between choice and prices, but not between buying and selling prices. Finally, simple heuristic rules can account for similarity effects on choice, but they are unable to account for attraction and compromise effects. In short, despite the increased flexibility provided by allowing mixtures of strategies, these models have not yet proven capable of providing a coherent explanation for many of the well-established findings.

Several computational models were presented, but two in particular stand out as most promising for meeting the challenge of this chapter. Both DFT and the leaky accumulator model provide coherent explanations for similarity, attraction, and compromise effects on choice. Furthermore, both of these models can predict how time pressure moderates these effects. In fact, the two models are based on very similar principles for making a choice, that is, a race between accumulators of preference to a threshold. The models differ in terms of their details concerning lateral inhibition and nonlinear accumulation. However, DFT has been applied more broadly than the leaky accumulator; the former also accounts for preference reversal among different measures of preference (choice vs. prices and buying prices vs. selling prices) as well as the paradoxes of risky decision making (Allais and stochastic dominance paradoxes). In conclusion, these two "accumulation to threshold" models provide explanatory power that goes beyond the algebraic utility models and the simple heuristic models.

Critics of computational models may claim that the power of these models comes at a cost of increased complexity. However, it is important to note that computational models may have the same number of (or even fewer) free parameters than the algebraic utility models applied to the same domain (see Section 3.4; cf. Johnson & Busemeyer, 2005a). By focusing on underlying cognitive processes, computational models can provide parsimonious explanations for broad collections of puzzling behavioral phenomena. In addition, computational models make precise predictions not possible with other approaches. Unlike typical utility models, computational models are dynamic and thus offer deliberation time predictions. Many of these models – including DFT, the focus of the current chapter – also account for variability in human behavior, in contrast to deterministic approaches, such as RDU theory and simple heuristic models.

6.2. *Connections to Computational Modeling in Judgment*

There are now a variety of computational models relevant to judgment and decision making research. Connectionist models of social reasoning are reviewed in Chapter 18 in this volume, and Stasser (2000) has considered computational models for information sharing in group decision making. Instance-based memory models of Bayesian inference (Dougherty, Gettys, & Ogden, 1999) and decision making (Stewart, Chater, & Brown, 2006) have been developed. Stochastic models of confidence judgments have been proposed (Brenner, Griffin, & Koehler, 2005; Erev, Wallsten & Budescu, 1994; Wallsten & Barton, 1982; Wallsten & Gonzalez-Vallejo, 1994). Several computational models of strategy learning have appeared (Busemeyer & Myung, 1992; Johnson & Busemeyer, 2005b; Rieskamp & Otto, 2006). This chapter is directed at decision making rather than reasoning or inference (but see Chapter 11 in this volume); it is focused on performance rather than memory or learning models; and it concerns individual as opposed to group decision processes.

7. Conclusion

This chapter discussed how a particular computational model could account for a wide variety of empirical trends that have resisted a coherent explanation by models cast in the dominant framework. This accomplishment was made possible by considering an alternative level of analysis, rather than attempting to further modify the utility framework. In addition, computational models have distinct advantages – both theoretical and practical – over contemporary approaches toward the study of decision making. Hopefully, more and more researchers will appreciate these advantages and contribute to an expanding and interesting literature involving computational models.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le riske: Critique des postulats et axiomes de l'ecole Americaine. *Econometrica*, *21*, 503–546.

Allais, M. (1979). The so-called Allais paradox and rational decisions under uncertainty. In O. Hagen & M. Allais (Eds.), *Expected utility hypotheses and the Allais paradox* (pp. 437–681). Dordrecht: Reidel.

Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, *44*, 310–329.

Belavkin, R. V. (2006). Towards a theory of decision-making without paradoxes. In D. Fum, F. D. Missier & A. Stocco (Eds.), *Proceedings of the Seventh International Conference on Cognitive Modeling* (pp. 38–43). Trieste, Italy: Edizioni: Goliardiche.

Bell, D. E., Raiffa, H., & Tversky, A. (1988). *Decision making: Descriptive, normative, and prescriptive interactions*. Cambridge, UK: Cambridge University Press.

Birnbaum, M. H. (2004). Causes of Allais common consequence paradoxes: An experimental dissection. *Journal of Mathematical Psychology*, *48*(2), 87–106.

Birnbaum, M. H., & Navarrete, J. B. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty*, *17*, 49–78.

Birnbaum, M. H., & Zimmermann, J. M. (1998). Buying and selling prices of investments: Configural weight model of interactions predicts violations of joint independence. *Organizational Behavior & Human Decision Processes*, *74*(2), 145–187.

Bostic, R., Herrnstein, R. J., & Luce, R. D. (1990). The effect on the preference-reversal phenomenon of using choice indifference. *Journal of Economic Behavior and Organization*, *13*, 193–212.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*, 409–432.

Brenner, L., Griffin, D., & Koehler, D. J. (2005). Modeling patterns of probability calibration with random support theory: Diagnosing case-based judgment. *Organizational Behavior and Human Decision Processes*, *97*(1), 64–81.

Busemeyer, J. R., & Diederich, A. (2002). Survey of decision field theory. *Mathematical Social Sciences*, *43*, 345–370.

Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology:General*, *121*, 177–194.

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations for decision field theory. *Mathematical Social Sciences*, *23*, 255–282.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459.

Dhar, R., Nowlis, S. M., & Sherman, S. J. (2000). Trying hard or hardly trying: An analysis of context effects in choice. *Journal of Consumer Psychology*, *9*, 189–200.

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology*, *41*, 260–274.

Diederich, A. (2003). MDFT account of decision making under time pressure. *Psychonomic Bulletin and Review*, *10*(1), 157–166.

Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *Journal of Mathematical Psychology*, *47*, 304–322.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory process model for judgments of likelihood. *Psychological Review*, *106*, 108–209.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review*, *69*, 109–135.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643–669.

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, *101*(3), 519–527.

Gigerenzer, G., Todd, P.M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, *69*, 623–638.

Grossberg, S., & Gutowski, W. E. (1987). Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review*, *94*, 300–318.

Guo, F. Y., & Holyoak, K. J. (2002). Understanding similarity in choice behavior: A connectionist model. In W. Gray & C. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual conference of the Cognitive Science Society* (pp. 393–398). Hillsdale, NJ: Lawrence Erlbaum.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, *128*, 3–31.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*, 90–98.

Huber, J., & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research*, *10*(1), 31–44.

Johnson, J. G., & Busemeyer, J. R. (2005a). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychological Review*, *112*(4), 841–861.

Johnson, J. G., & Busemeyer, J. R. (2005b). Rule-based Decision Field Theory: A dynamic computational model of transitions among decision-making strategies. In Betsch, T. & Haberstroh, S. (Eds.), *The routines of decision* *making* (pp. 3–20). Mahwah, NJ: Lawrence Erlbaum.

Johnson, J. G., & Busemeyer, J. R. (2006). Trading "as if" for "as is" models of cognition: A computational model of the attention processes used to generate decision weights in risky choice. *Cognitive Psychology*. Manuscript submitted for publication.

Kahneman, D., & Tversky, A. (1979). Prospect theory. *Econometrica*, *47*, 263–292.

Laming, D. R. (1968). *Information theory of choice reaction times*. New York: Academic Press.

Leland, H. E. (1998). Agency costs, risk management, and capital structure. *Journal of Finance*, *53*, 1213–1243.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*, 46–55.

Lindman, H. R. (1971). Inconsistent preferences among gambles. *Journal of Experimental Psychology*, *89*, 390–397.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, *40*, 77–105.

Loomes, G., Starmer, C., & Sugden, R. (1992). Are preferences monotonic? Testing some predictions of regret theory. *Economica*, *59*(233), 17–33.

Lopes, L. L., & Oden, G. C. (1999). The role of aspiration level in risky choice: A comparison of cumulative prospect theory and SP/A Theory. *Journal of Mathematical Psychology*, *43*, 286–313.

Luce, R. D. (2000). *Utility of gains and losses*. NJ: Lawrence Erlbaum.

Mellers, B. A., Schwartz, A., & Cooke, A. D. J. (1998). Judgment and decision making. *Annual Review of Psychology*, *49*, 447–477.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 226–300.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, *43*, 87–131.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organizations*, *3*, 323–343.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: A review of research on preferential choice. *Journal of Economic Literature*. *44* 631–636.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multi-alternative decision field theory: A dynamic connectionist model of decision-making. *Psychological Review*, *108*, 370–392.

Savage, L. J. (1954) *The foundations of statistics*. New York: Wiley.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science*, *15*, 331–336.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*, 158–174.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29*(3), 281–295.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*(5), 364–371.

Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review*, *73*, 596–605.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review*, *102*(3), 567–593.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, *38*, 332–382.

Stasser, G. (2000). Information distribution, participation, and group decision: Explorations with the DISCUSS and SPEAK models. In Computational modeling of behavior in organizations: The third scientific discipline (pp. 135–161). Washington, DC: American Psychological Association.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*.

Thagard, P., & Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In A. Ram & D. B. Leake (Eds.), *Goal-driven learning* (pp. 439–454). Cambridge, MA: MIT Press.

Tversky, A. (1972a). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281–299.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference dependent model. *Quarterly Journal of Economics*, *106*, 1039–1061.

Tversky, A.,& Kahneman, D. (1992). Advances in prospect theory: Commulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Tversky, A., & Sattath, S. (1979). Preference trees. *Psychological Review*, *86*, 542– 573.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*, 371–384.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *102*(3), 550–592.

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamic models of multi-alternative choice. *Psychological Review*, *111*, 757–769.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wallsten, T. S., & Barton, C. (1982). Processing probabilistic multidimensional information for decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 361–384.

Wallsten, T. S., & Gonzalez-Vallejo, C. (1994). Statement verification: A stochastic model of judgment and response. *Psychological Review*, *101*, 490–504.