

10 Parallel-Processing Mechanisms and Processing of Organized Information in Human Memory

Roger Ratcliff
Dartmouth College

The main aim of this chapter is to describe research in two areas of cognitive psychology that are relevant to the interests of workers in Artificial Intelligence (AI), cognitive science, and neural modeling. The chapter is divided into two parts: First, a parallel-processing associative model for recognition of independent items is presented. This model can be viewed as an up-to-date example of psychology's contribution to parallel-processing systems. The second part of the chapter presents some experimental results and psychological models concerned with processes and the representation of organized information in human memory.

In the area of neural modeling researchers are usually concerned with developing models that are able to account for behavioral results. Researchers in Artificial Intelligence have two slightly divergent aims: The first is to develop computer programs that perform certain tasks with humanlike or better intelligence. The second aim is to use any insights as theories of human performance and to model behavioral results. For both of these it is necessary to select aspects of human performance for comparison. The neural modeler or AI researcher has to find and select behavioral data that can provide good tests for his or her model but this process can be rather haphazard. For example, the selection may be made on the basis of the last few articles read, or it may be the latest paradigm of the local cognitive psychologist.

In the first section of this chapter I describe a parallel-processing associative model for recognition that accounts for a great deal of data in its domain of applicability and use this model as a case study in examining the relationship between theory and data. In particular several important properties of reaction time data are presented and several major properties of the human processing

system are brought into focus and described in terms of the model. It should be made clear at the outset that the model presented is composed of two parts: a mathematical model and a metaphor that is used to elucidate the model and present a particular view of the processing system. Several of the neural network models presented in this book may be quite compatible with the mathematical model, yet may present a different metaphor for the model (see Hinton, Chapter 6, this volume).

10.1. A MODEL FOR ITEM RECOGNITION

Reaction time has been used as the main dependent variable for developing and testing many models in cognitive psychology. I argue that the statistic, mean reaction time, is almost totally inadequate for such purposes (though it has proved useful in getting some areas of research started). One problem is that certain models of the recognition process, serial scanning models, can be mimicked by several other kinds of models, for example, parallel-processing models and direct-access models (based on strength theory), at the level of mean reaction time. Thus it is difficult to press any claims as to the nature of the processes involved in the tasks under study. Another problem that has arisen is that many models that are quite adequate at the level of mean reaction time have serious problems with the shape of reaction time distributions. Models of processing should account for the overall shape of reaction time distributions, that is, produce distributions that are positively skewed. If such care is not taken, then it is possible to produce models that are falsified by the data that they were designed to fit (see Ratcliff & Murdock, 1976).

Another factor of considerable importance is the relationship between accuracy and reaction time. There is always a relationship between accuracy and reaction time in recognition data: In a condition in which recognition is easy, a response is usually accurate and fast, whereas in a condition in which recognition is harder, a response is less accurate and slower. (This reflects difficulty of retrieval rather than the speed-accuracy tradeoff that will be discussed later.) It is relatively easy to produce a model that deals just with reaction time or just with accuracy, but I believe that it is important for any model of processing to deal with the relationship between accuracy and reaction time. A third factor of considerable importance is flexibility in the processing system. Subjects can choose to respond slower in most tasks in order to gain accuracy, or they can sacrifice accuracy to increase speed. Flexibility should therefore be central to any model of processing.

I now present a model for the process involved in recognizing whether an item is a member of a previously presented list. This model attempts to account for the shape of reaction time distributions, the relationship between speed and accuracy, and the flexibility in the processing system. In a typical procedure, a list of

items (words, letters, numbers, or pictures) is presented to a subject. This list is called the memory set. The subject is then presented with a test item and has to respond as to whether that item was in the list. Typically accuracy and reaction time are recorded and form the basic data.

According to the model the test item is encoded and then compared to the representation of each item in the memory set simultaneously (i.e., in parallel). Each individual comparison is accomplished by a random walk process. A decision is made when any one of the comparisons ends in a match or when all of the comparisons end in a nonmatch. When the decision has been made, the appropriate response is initiated. The overall scheme is shown in Fig. 10.1.

This scheme (random walk comparison process: comparisons carried out in parallel, self-terminating for positive responses and exhaustive for negative responses) provides the core of the mathematical model. It is later shown that this model does a good job in dealing both qualitatively and quantitatively with a relatively large amount of data from recognition studies. The metaphor that follows is not the only possible description of the processing system, but it does provide a reasonable way to understand the operation of the mathematical model. In a later section the relationship between the mathematical model and neural network models (providing a different metaphor for the model) is examined.

The metaphor used to describe the interaction between the test item and items in memory is a resonance metaphor. The test item causes memory items to resonate: the greater the resonance, the closer the match; the smaller the reso-

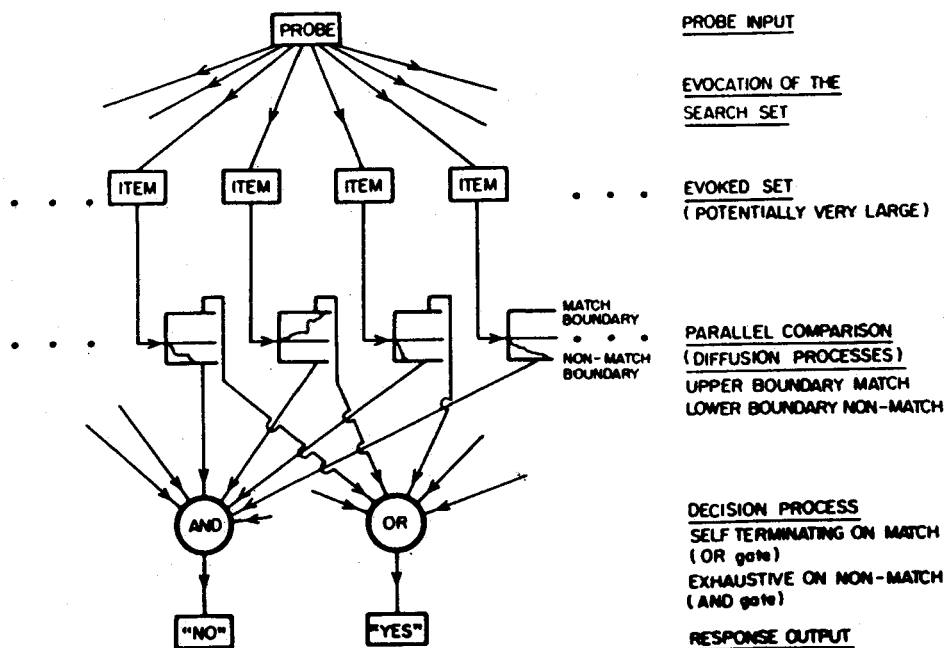


FIG. 10.1. An overview of the recognition model. Copyright 1978 by the American Psychological Association. Reprinted by permission.

nance, the poorer the match. The resonance metaphor is used to indicate that items outside the immediately preceding memory set are accessed in the comparison process. Atkinson, Herrmann, and Wescourt (1974) have shown that items in a set of instructions enter the comparison process. They presented items from the instructions as negative items in the test list in a recognition experiment. Reaction time to these items was slower than reaction time to control items. Monsell (1978) has traced out the decay function and found that a test item last presented several lists back can influence speed and accuracy of a "no" response. The size or amount of resonance drives the random walk: the greater the resonance, the greater the rate of approach to the match boundary; the smaller the resonance, the faster the approach to the nonmatch boundary.

The size of the resonance is determined by several factors. First, it appears that every kind of similarity enters the comparison; for example, Juola, Fischler, Wood, and Atkinson (1971) performed an experiment with words as study and test items. They looked at three types of negative test items (items to which the correct answer was "no"): homophones and synonyms of stimulus words and neutral words. They found that synonyms were 60 msec slower than neutral words and that homophones were 120 msec slower than neutral words. When homophones were broken down into two classes, visually similar and visually dissimilar, increases in reaction times were 200 msec and 40 msec, respectively. These results show that semantic, visual, and acoustic similarities enter the comparison between a test item and stimulus items. Morin, DeRosa, and Stultz (1967) have shown that the more numerically remote a negative probe is from the memory set, the faster the response (e.g., a negative response to a 1 as a test item is faster than a negative response to a 6 if the study items were 7, 8, and 9). Besides similarity, how recently a study item was presented affects both reaction time and accuracy in many different paradigms.

From this discussion it can be seen that many different kinds of information enter the comparison process. The probe information interacts with memory trace information producing a yes or no response. In order to maintain the notion that the memory trace consists of many different kinds of information, the term relatedness is used to describe the overall amount of match. An example of the way relatedness may be derived independently is given in Rips, Shoben, and Smith (1973). They obtained ratings from subjects for the similarity between pairs of birds and pairs of animals. From these similarity ratings multidimensional scaling solutions were obtained for birds and animals separately. Two-dimensional solutions were adequate; the dimensions were identified as predacity and size. To determine the relatedness between two concepts, the euclidean distance between the two concepts in the two-dimensional space can be assessed. This single-measure relatedness is used as the value of drift that drives the random walk process to determine whether probe and memory items match. If relatedness is high, then drift toward the match boundary is rapid; if relatedness

is low, then drift is towards the nonmatch boundary. In the mathematical model relatedness has variability in that two nominally equivalent items (e.g., having the same serial position in the memory set) may have different relatedness values, for example, because they have been learned to different levels.

The comparison process can be illustrated by supposing that the probe and memory items are represented by a vector of features. The comparison then proceeds by a gradual accumulation of feature matches (which could be either a serial or parallel process). Each time a match occurs between a feature in the probe and a feature in the memory item a step is taken toward the match boundary in the random walk. Each time a nonmatch occurs, a step is taken toward the nonmatch boundary. Relatedness represents the overall number of feature matches between the probe and the memory item. In terms of the resonance metaphor the size of the resonance represents the average rate of feature matches.

One source of variability, variability in relatedness, has already been mentioned. It can be seen that there is a second source of variability and that is variability in the rate of accumulation of evidence. Two probe vectors, with the same number of feature matches to a particular memory-item vector, may have differences in the order of matches and nonmatches so that one comparison may have mainly feature matches in the first part of the comparison, and another comparison may have mainly nonmatches in the first part of the comparison. These two processes would have quite different times to reach the match or nonmatch boundary (the processes may even reach different boundaries). Thus we can identify two different sources of variability in the model; variability in relatedness and variability in the comparison process.

In the specific mathematical model, the comparison process is modeled by the continuous analog of the random walk: the diffusion process. The diffusion process is the component of the model that accounts for such factors as reaction time distributions, speed-accuracy relationships, and variable criteria, that is, flexibility in processing. The critical assumption made is that relatedness is proportional to the average drift in the diffusion process. Relatedness between a probe and a memory-set item when the probe matches the memory item is assumed to be distributed normally with mean u and variance η^2 and for a nonmatching probe and memory item, relatedness is again assumed to be distributed normally with mean v ($u > v$ usually) and variance η^2 . A criterion is set between u and v such that values of drift on the u side of the criterion are positive (drift towards the match boundary) and values of drift on the v side of the criterion are negative. Figure 10.2 illustrates the relatedness distributions, the random walk process, and the diffusion process.

Reaction time distributions of the usual empirical shape are produced by the diffusion model. Figure 10.3 shows the way in which the normal distribution maps into a positively skewed reaction time distribution. In addition variability in drift works to make the distributions more positively skewed. As can be seen

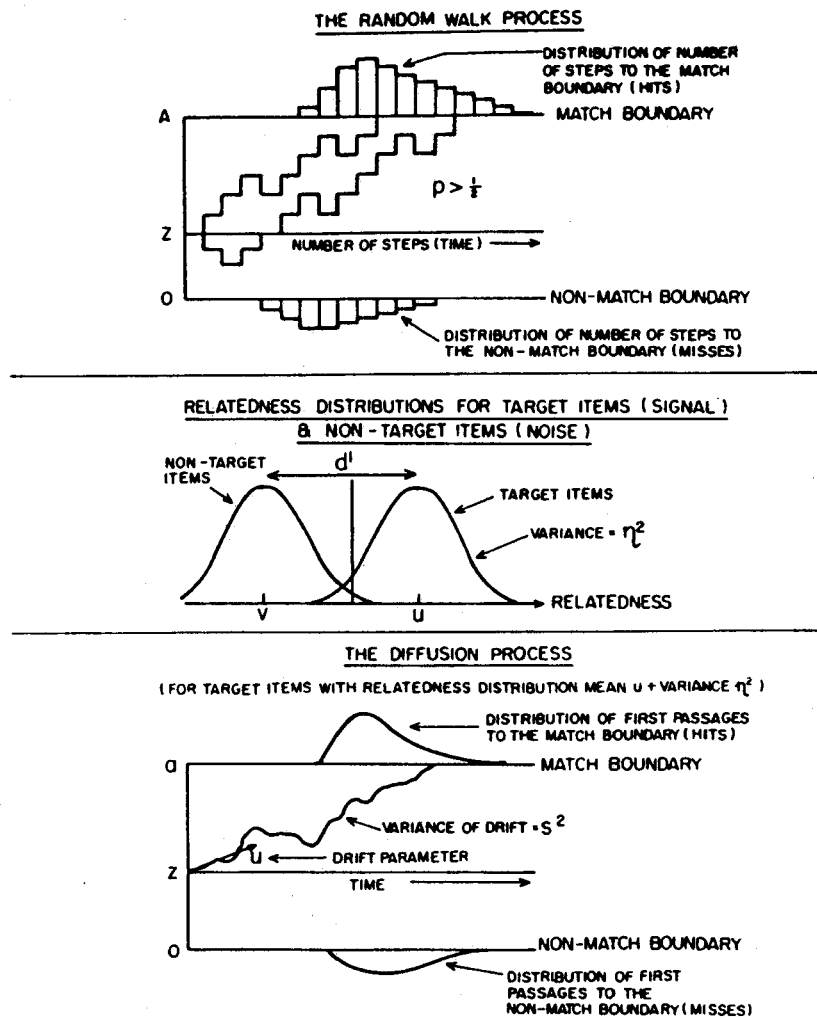


FIG. 10.2. An illustration of the random walk and diffusion process together with relatedness distributions that drive the diffusion process. Copyright 1978 by the American Psychological Association. Reprinted by permission.

from Fig. 10.3, a prediction of the model is that as relatedness decreases, the mean and mode of the distribution both increase and diverge as is seen in reaction time data.

There are three variable criteria in the model: the zero point of relatedness, and the positions of the two boundaries. It is assumed that all these criteria are to some extent under the subject's control. The way subjects are able to control their speed-accuracy criteria is by adjusting the position of the match and nonmatch boundaries. The criterion in relatedness can be adjusted, in the same way that the criterion in signal detection theory can be adjusted, to vary the relative numbers of false positive responses and false negative responses. These two sets of criterion adjustments are not entirely independent in that adjustments in any of the three criteria produces changes in both speed and accuracy of responses.

From Fig. 10.2, it can be seen that the random walk process integrates both reaction time and accuracy in a single theoretical mechanism. This integration allows speed-accuracy trade-off to be explained in terms of changes in boundary positions. The closer the match and nonmatch boundaries are to the starting point of the random walk, the faster and less accurate are responses; the further away from the starting point, the slower and more accurate are the responses (with relatedness values and relatedness criterion held constant). There are two major ways in which speed-accuracy trade-off may be studied. First subjects can be induced to respond with either speed or accuracy by instructions. This mode of responding can be termed information-controlled processing; the subject determines when to respond based on the amount of information accumulated. The model as described so far is concerned with this mode of responding. Second, the experimenter may determine the time at which the subject will respond using a deadline or signal to respond. In this case, the mode of processing may be termed time-controlled processing. Figure 10.4 shows the way the distribution of evidence spreads in time-controlled processing. Initially the evidence begins at the

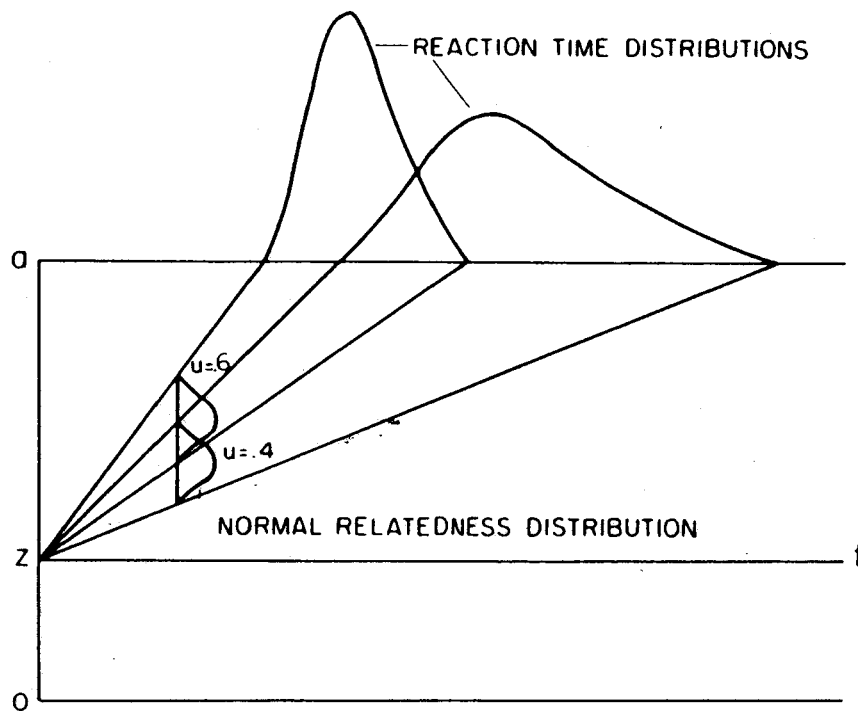


FIG. 10.3. A geometrical illustration of the mapping from a normal relatedness distribution to a skewed reaction time distribution (with variance in drift $s^2 = 0$). (Note that as relatedness decreases, the distribution tail skews out. a represents the distance between the bottom and top boundaries of the diffusion process; z represents the distance between the bottom boundary and the starting point; and u represents the mean of the normal relatedness distribution.) Copyright 1978 by the American Psychological Association. Reprinted by permission.

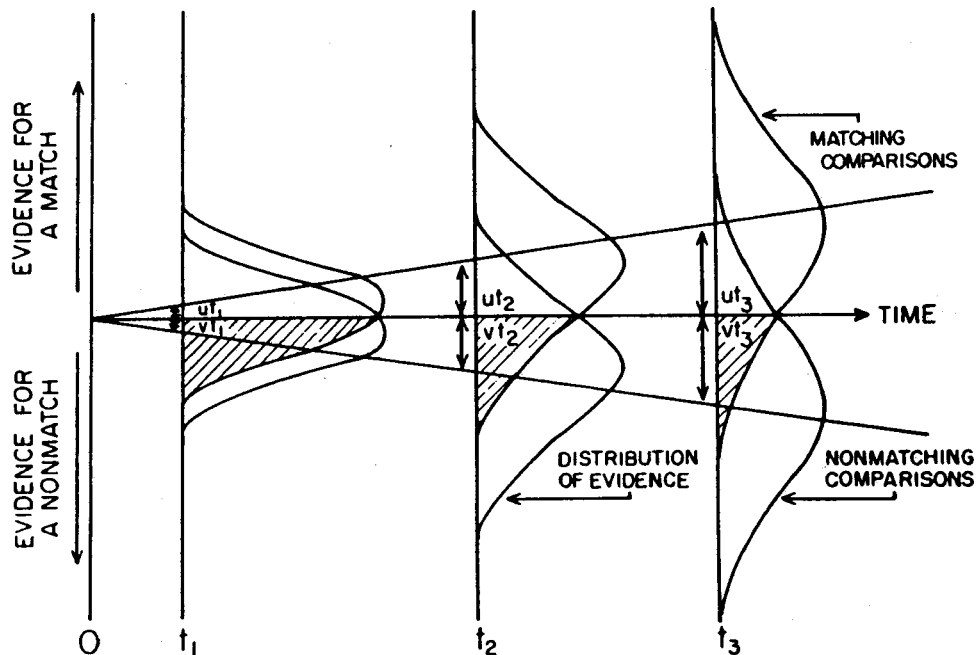


FIG. 10.4. Spread of evidence as a function of time in the unrestricted diffusion process for one matching and one nonmatching process. (At time t_1 , there is a large amount of overlap; at time t_3 , the overlap has reached asymptote, with asymptotic $d' = (u - v) / \eta$ [See middle panel of Fig. 10.2]. u is the mean of the match relatedness distribution, and v is the mean of the nonmatch relatedness distribution.) Copyright 1978 by the American Psychological Association. Reprinted by permission.

starting point z ; as processing proceeds, evidence spreads out with the variance a function of time (t). At large values of time the distributions of evidence (both for matching and nonmatching comparisons) will tend to an asymptotic form: All comparisons with positive relatedness will have reached a position on the positive side of the starting point, and comparisons with negative relatedness will have drifted negative. Thus it can be seen that accuracy will asymptote as a function of time at the d' value defined by the relatedness values (and this is observed in practice; see Reed, 1976).

From this discussion it can be seen that the random walk comparison process is capable of accounting for the shape of reaction time distributions, the relationship between reaction time and accuracy, and the flexibility of processing evident in speed-accuracy trade-off.

Error Analysis and the Time Course of Evidence Accumulation

If it is possible to investigate the patterns of errors at different signal lags in a deadline or response signal experiment, then it is possible to examine the kind of information that is being used by the subject at different points during the time

course of accumulation of evidence. Pachella, Smith, and Stanovich (1978) performed experiments in which subjects were presented visually with one of four stimuli, the letters *B*, *C*, *D*, and *E* were required to name the stimulus at certain deadlines. Pachella et al. performed error analyses at each of the deadlines and found that the error patterns could be well fitted by an informed guessing model. The informed guessing model in this application had the following states: *BCDE*, *BD*, and *CE* (determined by fits of the model to the data; other sets were shown to be unimportant). If subjects had no information they guessed from the set *BCDE*; if they had some information, then they either had information that distinguished the set *BD* from *CE* or total information. As the deadline became longer, the probability of being in one of the confusion sets became smaller. From the results it can be seen that it is possible to trace out the time course of accumulation of evidence and to find out the form of the evidence being accumulated at particular points during the time course.

McClelland (1979) has presented a model for the case where processes operate in cascade; that is, one processing stage can make use of the partial results of a previous processing stage. It is possible to account for the results of Pachella et al. (1978) in terms of this cascade model by supposing that at short lags the response is made on the basis of information from early processing stages in which the subject has little information (i.e., is in a guessing state). At intermediate lags the response is made on the basis of partial information from which the subject can distinguish sets *BD* from *CE*, and at longer lags the subject uses information from later processing stages in which total information is available. Ratcliff (1980) has presented the mathematics necessary to deal with the diffusion process when the rate of information accumulation varies as a function of time. In this case, experiments of the kind performed by Pachella et al. (1978) will provide information that may help identify the processing stages in a cascade-type model or that might identify the type of evidence being accumulated at a particular time in a diffusion model.

The Decision Process

The decision process is conceived of as a process in which the results of many comparisons are combined to produce a single yes/no decision. The process terminates when one comparison ends in a match leading to the production of a positive decision but has to wait for all comparisons to terminate in a nonmatch to produce a negative response.

It may be somewhat difficult to see how "yes" and "no" responses can have about the same reaction time, as is observed experimentally, when negative responses require that all comparisons terminate whereas positive responses require only one comparison to terminate with a match. The model allows the relative speed of yes and no responses to be determined by the relative starting point to boundary distances, and in fits of the model it is found that the starting point to match boundary distance is in general greater than the starting point to

nonmatch boundary distance. Because subjects are able to manipulate these boundary positions (in the formulation of the model), they can vary the relative speeds of match and nonmatch processes that lead to yes and no responses (note that accuracy will also covary). A direct analogy can be found in the analysis of detection data and signal detection theory. In signal detection theory the proportions of correct yes and correct no responses are used to compute two measures of performance, a measure of discriminability and a measure of the criterion that subjects set. The criterion setting reflects the relative certainty subjects place on yes and no responses: Subjects can choose to respond "yes" only when very sure thus producing relatively few but very accurate yes responses, or subjects can choose to respond "no" only when very sure, or at any point in between. Ratcliff and Hacker in an unpublished experiment investigated criterion effects on reaction time by measuring reaction time in two experiments in which subjects were encouraged to be sure when responding "yes" in one condition and to be sure when responding "no" in another condition. In a recognition memory task reaction time was found to covary with accuracy in that when subjects were responding in the sure yes-condition, accuracy for yes responses was high and reaction time for yes responses was 168 msec slower than reaction time for no responses. In the sure no-condition accuracy for no responses was high and reaction time for no responses was 203 msec longer than for yes responses. These data can be modeled by assuming that the random walk boundaries are adjusted so that in the sure yes-condition the match boundary is far away from the starting point and the nonmatch boundary is relatively close to the starting point. This adjustment will lead to the observed pattern of data. The main point to note is that the relative speed of yes and no responses is under the control of the subject and is another indication of the flexibility in processing discussed earlier.

Summary of the Model

The model I present deals with aspects of experimental data that are held to be critical for any reasonable model of item recognition. The relationship between accuracy and reaction time is implicit in the random walk comparison process and the random walk process guarantees reaction time distributions of the correct shape. Flexibility in processing is allowed by the variable (subject-controlled) criteria, and adjustment of these criteria allow the modeling of such things as speed-accuracy trade-off. As well as accounting for the experimental data within any single paradigm, the model accounts for performance on several different paradigms and thus allows comparisons to be made among those paradigms (see Ratcliff, 1978, for details).

Comparison with Anderson's Model

It is easy to see relationships between the above model and certain neural network models. In particular, the model can be related to a recognition model

developed by J. A. Anderson (1973). In that model the memory representation of a set of items is a vector of N elements. A particular item is represented by a specific vector of N elements. The memory representation of several traces is then the vector sum of the individual traces. The match between a probe and the overall memory is simply the dot product between the probe vector and the overall memory vector (i.e., each element in the probe vector is multiplied by the corresponding element in the memory vector). In Anderson's model positive and negative evidence (element matches and nonmatches, respectively) are accumulated separately until one of them exceeds a fixed criterion. If this were modified so that positive and negative evidence canceled, then the comparison process would be a random walk. The model suffers the problem, however, that there is no separate record of each memorized item. Thus, on the basis of this memory, subjects would be unable to judge such things as frequency of occurrence of items in a list and would be unable to determine in which list an item had been presented (these arguments are the same as those presented against strength theory, see Anderson, J.R., & Bower, 1972; Wells, 1974).

In contrast to neural network models of the type developed by Anderson, the retrieval model I describe maintains that the representations of items in memory are functionally independent; that is, there are separate representations of the occurrence of individual items. It turns out that in fitting response signal data (Reed, 1976), a model that assumes that all information about the study material is combined in a single vector would not produce adequate fits. The inadequacy of such a fit is one reason that the item recognition model described above maintains separate representations for each item encoded.

General Discussion

For neural modelers whose area of interest is higher cognitive functioning, one of the major aims of modeling is to realize psychological functions such as recognition, association, and categorization in terms of reasonable neural models. There are presently available models that do a good job of representing similarity, partial matching of items, reconstruction of stimuli from degraded probes, recognition, association, and categorization (see Anderson & Mozer, Chapter 7, and Kohonen et. al., Chapter 4, this volume). Each of the models does a good job of mimicking human performance within its domain and has desirable properties from the standpoint of system design.

There are three major problems with these kinds of models. The first is that they seldom add much to our psychological understanding of the structure and processes that they model in that they rarely make strong predictions about how humans will perform in other tasks, they rarely integrate experimental paradigms so as to provide parameter invariance across tasks. The second problem concerns mimicking of the different neural schemes by each other and the difficulty of separating these schemes on the basis of desirable properties or fits to data. The third problem concerns the completeness of such models; each of the models can

be considered a building block or an element in an intelligent system, but there seems to be no attempt to develop a control structure that fits these elementary processes together to produce an intelligent system. There are also several more detailed difficulties, and these include problems in modeling semantic networks with labeled relationships that appear to be a prerequisite to representing world knowledge and allowing the system to make inferences (see Hinton's chapter in this volume).

Artificial Intelligence models seem to be complementary in some respects. The main concern is with a whole system (no matter how modest) that in fact performs the tasks it was designed to perform. The notion of control and of fitting elementary processes together is of central importance. In these models (see Fahlman, Chapter 5, this volume) world knowledge and relational information are used in the knowledge base, and such information allows inferences to be made. However, such Artificial Intelligence models have difficulty in representing similarity (except by relative distance in the network) and in performing partial matching (matching an incomplete probe against memory). It is also difficult to see how such models would be implemented in the nervous system.

Both of these areas of research have insights to gain from psychology. At some point a researcher has to choose some set of phenomena to model. If the choice is made on the basis of what an intelligent system should do, then this often boils down to the use of informal psychological data. Psychology can point to data that may rule out various alternate, intuitively appealing models. For neural modelers, psychological results can often form the basis for the whole modeling effort, and for Artificial Intelligence researchers, psychological results can point to more useful methods of organizing the theoretical processing system. It should be noted that psychology often gains insights from Artificial Intelligence, and some theories developed first in Artificial Intelligence are then taken as psychological theories and subjected to experimental test.

In the next section I present a brief discussion of some psychological research on the topic of the organization of information in memory and processes involved in encoding and retrieval.

10.2. ORGANIZED MEMORY: PROCESS AND STRUCTURE

There seem to have been relatively few attempts by cognitive psychologists to communicate the present certainty of phenomena and theories to groups such as Artificial Intelligence researchers and neural modelers. This section is a small attempt to present some of the more recent findings in the general area of research concerned with the structure of semantic knowledge and the representation of text in memory and with processes involved in manipulating this information. In particular I discuss topics such as the distinction between automatic and

strategic processes, a line of research that may allow us to decide whether inferences are made at input or output in text processing and the processing of world knowledge (verification of well-known facts).

Automatic and Strategic Processes

The processes involved in priming have come under considerable scrutiny recently. The main result is that presenting one concept (word or letter) can activate another concept leading to a faster match, lexical decision, or recognition for the second concept. For example, in a letter-matching task (Posner & Snyder, 1975), the subject is to respond as to whether two target letters presented simultaneously are the same or different. The same response is speeded if a prime letter, presented before the target letters, matches the target letters (i.e., a facilitation effect).

Automatic and strategic components of priming can be separated by studying the time course of this facilitation, using a procedure in which the target letters are presented at a variable interval (e.g., at 50, 100, 200, and 400 msec) after the prime. The priming effect appears to have two components. The first component has been called an automatic component; it is characterized by a very rapid onset; in many experiments, prime onset to target onset as short as 100 msec is sufficient to produce facilitation. The second component has been termed strategic facilitation and is characterized by a much less rapid onset, often several hundred milliseconds.

Automatic and strategic processes can also be distinguished by probability manipulations. Automatic processing has no inhibition associated with low probability alternatives whereas strategic processing often has inhibition associated with low probability alternatives. For example, in the letter-matching paradigm, if the prime does not match the test letters (in a condition in which there is a low probability of a prime-target nonmatch) then there is inhibition; that is, the same response is slowed (Posner, 1978, p.100).

Neely (1977) provided a particularly clear example of the separation of automatic and strategic components of facilitation in a lexical decision task. Subjects were presented with a prime word, to which they made no response, followed by a target. The subjects were required to decide whether the target letter string was a word. The prime was one of three category names or a row of xs. Subjects were told that if the prime was *bird* then they should expect the target letter string to be a member of the bird category if the letter string were a word. If the prime was *building* then the target word would be a body part with high probability. If the prime was *body* then the target letter string would be a building part with high probability. Subjects were explicitly told to shift their attention when they saw the *building* or *body* prime to the expected category. The time course of processing was examined by varying the prime to target delay. For the categories *building* and *body* there was facilitation for a target that

was a member of the prime category at short delays (250 msec). This was interpreted as automatic activation. At longer delays (750 msec) the response to a target that was a member of the prime category in this shift condition was inhibited (a longer reaction time). The expected target (a body part if *building* was the prime) produced facilitation at longer delays but not at short delays. These results were interpreted as supporting the automatic-strategic distinction.

A similar series of questions can be asked about priming in recognition. McKoon and Ratcliff (1980) and Ratcliff and McKoon (1978) have used priming in recognition as a technique for examining the structure of text in memory. For example, Ratcliff and McKoon (1978) presented sentences to subjects for study and then tested single words for recognition. Response time to a word immediately preceded in the test list by a word from the same sentence was 100 msec faster than the response time to a word immediately preceded by a word from a different sentence. The size of the priming effect was found to be greater if the priming pair were from the same proposition than if the priming pair were from different propositions. This result was taken as support for the view that the structure of sentences is propositional. McKoon and Ratcliff (1980) showed that the magnitude of the priming effect varies as a function of the distance between the prime and test words in propositional structure of the paragraph. Thus the technique provides an index of the distance between two propositions in terms of the size of the priming effect. The priming effect in recognition shows that activation is not just reserved for preexisting semantic networks. Accessing a concept that was just studied in text serves to activate concepts related in the text. The amount of activation varies as a function of the relative distance between the concepts in the text. The question arises as to whether the priming effect is automatic or strategic. Ratcliff and McKoon (in press, a) have investigated automatic and strategic components of priming in recognition. In the first experiment it was found that the probability that a priming pair occurred in the test list had no effect on the size of the priming effect. This suggests that the priming effect is automatic in the sense of Tweedy, Lapinski, and Schvaneveldt (1977). In the second experiment the time course of processing was examined. A prime was presented to which the subject was not required to respond. Following this by a variable amount of time (e.g., either 50, 150, 450, 850 msec) the test word was presented for recognition. It was found that facilitation (when the prime and test word were from the same sentence) had been produced by 150 msec. Inhibition (when the prime and test words were from different sentences) occurred later in processing and showed up by 450 msec. The third experiment was designed to investigate strategic priming. Subjects were presented with two sentences and told that if the prime word was from one sentence, then the test word would be from the other sentence with high probability, and they should attempt to switch to that sentence. It was found that it took somewhat longer than 750 msec for subjects to switch—there was little facilitation at 750 msec for words from different sentences, large facilitation at 1800 msec, but

large inhibition when the prime and test words were from the same sentence at 750 msec—which indicates that strategic priming in recognition in this paradigm takes considerably longer than automatic processing. Thus we can conclude that the priming effects reported by McKoon and Ratcliff as an index of paragraph structure are automatic priming effects because all intertest intervals were kept less than 200 msec.

The automatic–strategic distinction has implications for the earlier discussion about flexibility of processing. Strategic processes are those subject to flexibility such as changing criteria or selecting among alternative succeeding processes during the course of processing. Automatic processes are those processes that run off no matter what the subject attempts to do strategically. The distinction between automatic and strategic processes has considerable importance for models of human processing (see also Schneider & Shiffrin, 1977 for discussion of the development of automaticity) but as yet has been largely ignored in the areas of Artificial Intelligence and neural modeling. The distinction may prove of help in determining whether a particular kind of inference is made at the time of reading a text or at the time of retrieval of that text.

Semantic Verification Experiments

Recent psychological investigation into the structure of semantic memory originated when Collins and Quillian (1969) performed several experiments to test Quillian's network theory of semantic memory. In this theory, concepts such as *robin*, *bird*, *animal*, and *thing* are stored in a hierarchy with *thing* as the root node and other concepts branching off (e.g., *animal* would be one link from *thing*, *bird* one link from *animal*, and *robin* one link from *bird*). In the experiments, subjects were asked to verify statements such as "a robin is a bird," or "a robin is an animal." The prediction made was that the time required to verify the sentence is a linear function of the distance between the concepts in the memory representation. This prediction was verified and further experiments followed, but several problems were found in attempting to fit data from negative responses. Rips, et al. (1973) developed an alternative model of the representation of semantic information. Their model represents similarity in terms of overlap of semantic features instead of distance in terms of number of links in a network model. See Smith (1978) for a discussion of further properties of feature and network models. Rips et al. performed several experiments in which the variable semantic relatedness was controlled and varied, and they found that semantic relatedness was a much better predictor of reaction time than hierarchical distance (in fact it was even suggested that the variable hierarchical distance had no effect on reaction time). Smith, Shoben, and Rips (1974) developed a feature-matching model that accounted for reaction time differences in terms of feature overlap, where feature overlap was used to represent semantic relatedness. In addition, Smith et al. added a decision mechanism (similar to signal detection

theory) that was based on a model for item recognition developed by Atkinson and Juola (1973). This retrieval model predicted that the more related (the greater the feature overlap) are two concepts, the faster and more accurate the positive response, and the slower and less accurate the negative response. Thus, verifying "a robin is a bird" is faster than verifying "a penguin is a bird" and responding negatively to "a bird is a robin" is slower and less accurate than responding negatively to "a bird is a penguin." This prediction is upheld when the relationship tested is of the form category and member, but when the relationship is an antonym relationship, the prediction is contradicted by data (e.g., Glass, Holyoak, & Kiger, 1979). Antonym relationships are verified quite quickly, the more related the terms, the faster. For example, "is a brother a sister?" is responded to negatively more quickly than a more indirect (and less related) antonym such as "is a brother a female?" Holyoak and Glass (1974) have presented data that suggest that production frequency is a better predictor of reaction time in such semantic verification tasks. From their results they developed a model to account for negative decisions that involved production or search then checking. Lorch (1978) has presented data that suggests that the two separate factors, production frequency and semantic relatedness, both have effects on verification reaction time and accuracy.

In order to account for many of the problems found in the original Collins and Quillian model, Collins and Loftus (1975) presented a revision of the model. The model assumes that concepts are stored in a network with the links between concepts labeled (as before) and weighted (i.e., by weights that denote strength of association between the concepts). The mechanism for retrieval consists of two stages: First, activation spreads from concepts represented in the question to activate a portion of semantic memory; second, that active portion is evaluated. This model has two main problems. First, the spreading activation process is probably not able to produce the levels of successful retrieval that humans produce (see Anderson & Hinton, Chapter 1, this volume). Furthermore, Ratcliff and McKoon (in press, b) have shown that the time required for activation to spread through a semantic network is very small, activation spreads too fast to account for any temporal variability in reaction time data (i.e., reaction time differences between conditions). Thus it seems that the spreading activation component of this model is largely unnecessary. Second, the evaluation process is not spelled out in sufficient detail though several different matching processes are described, for example, using counterexamples or distinguishing properties. However it is quite unclear which combination of mechanisms are used to explain experimental data and how all these mechanisms may be coordinated in a processing system.

From this discussion it can be seen that the theoretical interpretation of semantic verification is no longer simple (as it was with the models of Collins & Quillian, 1969, or Rips et al., 1973). We can identify two important variables

that affect performance, semantic similarity and production frequency, and we can note that antonyms appear to be processed differently to category-member relationships. But as yet we have no general relatively complete model of semantic verification. It is my guess that no simple elegant model will be developed for the semantic verification task; rather the kind of model developed will incorporate several autonomous subprocesses as in the Collins and Loftus (1975) model.

10.3. CONCLUSIONS

A main concern of this chapter is the relationship between psychological data and theory. In the first part of the chapter, a parallel-processing, associative model for recognition is presented. Besides providing a recent example of psychology's contribution to the theme of this book, the model provides an example of a theory that is general (applies across a range of paradigms) yet also explains and fits data within its domain in considerable detail. At the core of the model is a random walk comparison process that relates accuracy and reaction time, accounts for the shape of reaction time distributions, and allows the flexibility in processing necessary to account for speed-accuracy trade-off and other criterion adjustments. A model such as this can be seen as a replacement for informal data because the model summarizes a great deal of data in its domain so that other modelers need only concern themselves with the predictions of the model as a first step in further development. In a great number of enterprises in cognitive psychology, theories of the generality and detail of the theory presented in the first part of this chapter are not available. The second part of the chapter describes recent developments in two areas of research, each of which is concerned with some aspects of the structure of semantic information and text in memory and the processes involved in encoding, accessing, and retrieving such information. First it is argued that an important characteristic of the human processing system is the distinction between automatic processes that have rapid onset and are not subject to flexibility of processing and strategic processes that have much slower onset and can be adjusted by subjects particular processing strategies. This distinction has not yet entered the areas of neural modeling and Artificial Intelligence. Second an empirical method of studying the structure and process of permanent knowledge, the study of semantic verification, is reviewed. At present it seems that explanations of semantic verification results can no longer be simple but that we can identify two important variables, semantic relatedness and production frequency, and it seems that antonyms are processed differently from category-member relationships. There is a great deal of psychological research into parallel processing and associative memory, and it is hoped that this discussion will prove useful to neural modelers, Artificial Intelligence researchers, and perhaps even psychologists.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by NICHD grant HD 13318-01. Figures were reproduced by permission of the American Psychological Association.

I would like to thank Gary Dell, Gail McKoon, Ben Murdock, and Ed Smith for comments on an earlier draft of this chapter.

REFERENCES

- Anderson, J. A. A theory for the recognition of items from short memorized lists. *Psychological Review*, 1973, 80, 417-438.
- Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review*, 1972, 79, 97-123.
- Atkinson, R. C., Herrmann, D. J., & Wescourt, K. T. Search processes in recognition memory. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
- Atkinson, R. C., & Juola, J. F. Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance (IV)*. New York: Academic Press, 1973.
- Collins, A. M., & Loftus, E. F. A spreading-activation theory of semantic processing. *Psychological Review*, 1975, 82, 407-428.
- Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 240-247.
- Glass, A. L., Holyoak, K. J., & Kiger, J. I. Role of antonym relations in semantic judgments. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, 5, 598-606.
- Holyoak, K. J., & Glass, A. L. The role of contradictions and counterexamples in the rejection of false sentences. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 215-239.
- Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. Recognition time for information stored in long-term memory. *Perception & Psychophysics*, 1971, 10, 8-14.
- Lorch, R. F., Jr. The role of two types of semantic information in the processing of false sentences. *Journal of Verbal Learning and Verbal Behavior*, 1978, 17, 523-537.
- McClelland, J. L. On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 1979, 86, 287-330.
- McKoon, G., & Ratcliff, R. Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior*, 1980, 19, 369-386.
- Monsell, S. Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, 1978, 10, 465-501.
- Morin, R. E., DeRosa, D. V., & Stultz, V. Recognition memory and reaction time. In A. F. Sanders (Ed.), *Attention and performance I*. Amsterdam: North Holland, 1967.
- Neely, J. H. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, 1977, 106, 226-254.
- Pachella, R. G., Smith, J. E. K., & Stanovich, K. E. Qualitative error analysis and speeded classification. In N. J. Castellan, Jr. & F. Restle (Eds.), *Cognitive theory, Volume 3*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.
- Posner, M. I. *Chronometric explorations of the mind*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.
- Posner, M. I., & Snyder, C. R. Attention and cognitive control. In Solso, R. L. (Ed.), *Information processing and cognition*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.
- Ratcliff, R. A theory of memory retrieval. *Psychological Review*, 1978, 85, 59-108.

- Ratcliff, R. A note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology*, 1980, 21, 178-184.
- Ratcliff, R., & McKoon, G. Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1978, 17, 403-417.
- Ratcliff, R., & McKoon, G. Automatic and strategic priming in recognition. *Journal of Verbal Learning and Verbal Behavior*, in press. (a).
- Ratcliff, R., & McKoon, G. Does activation really spread? *Psychological Review*, in press. (b).
- Ratcliff, R., & Murdock, B. B., Jr. Retrieval processes in recognition memory. *Psychological Review*, 1976, 83, 190-214.
- Reed, A. V. List length and the time course of recognition in immediate memory. *Memory & Cognition*, 1976, 4, 16-30.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1-20.
- Schneider, W., & Shiffrin, R. M. Controlled and automatic human information processing. *Psychological Review*, 1977, 86, 1-66.
- Smith, E. E. Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6). Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.
- Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 1974, 81, 214-241.
- Tweedy, J. R., Lapinski, R. H., & Schvaneveldt, R. W. Semantic-context effects on word recognition: Influence of varying the proportion of items presented in an appropriate context. *Memory & Cognition*, 1977, 5, 84-89.
- Wells, J. E. Strength theory and judgments of recency and frequency. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 378-392.