

Network Strength Theory of Storage and Retrieval Dynamics

Wayne A. Wickelgren
University of Oregon

The notion of strength is defined in several alternative ways for chains of associations connected in series and in parallel. Network strength theory is extended to handle retrieval dynamics for a network of associations, in a manner that permits various degrees of serial versus parallel processing through a chain of serially ordered associations. In the parallel-processing version, a speedy activation pulse passes through the chain and initiates a relatively slower retrieval process virtually simultaneously at each link. It is demonstrated that under many conditions, the theory yields the same storage and retrieval dynamics for a network as for any component association. The theory is applied to recall and recognition, semantic memory, speech recognition, and reading.

VERTICAL ASSOCIATIONS

Classical association theory has generally assumed, either explicitly or implicitly, that the association of two ideas A and B is by means of direct "horizontal" associations between A and B. A variety of considerations now make it appear very unlikely that human conceptual memory uses only direct horizontal associations of ideas.

First, the notion that learning a pair consists of strengthening direct forward and backward associations between the two elements does not appear to generalize satisfactorily to semantic memory for propositional material (e.g., Anderson & Bower, 1973; Rumelhart, Lindsay, & Norman, 1972; Schank, 1973). All of the component words or concepts of a sentence are not directly associated to each other, nor are the primary associations between temporally or spatially adjacent words or concepts (Anderson & Bower, 1973). Rather, as suggested by structural linguistic intuition (Chomsky, 1957), semantic memory for propositional materials appears to have an organized hierarchical (more generally, network) character, with

complex propositions having elementary propositional constituents, elementary propositions having phrasal constituents, and phrases having smaller phrasal constituents and ultimately concept constituents. Besides the structural linguistic evidence for this hierarchical organization (e.g., Chomsky, 1957), there is also considerable psycholinguistic evidence for this in studies of sentence memory and perception (see Anderson & Bower, 1973).

Second, even the memory for an elementary A-B paired associate does not generally appear to be the result of strengthening a direct horizontal connection from the internal representative of the A element to the internal representative of the B element of the pair. It is now perfectly clear that direct rote association of two members of a pair, far from being the fundamental building block of learning, is actually the hardest procedure for learning a pair. An enormous number of studies indicate that paired-associate learning is greatly facilitated by (a) embedding a word pair into a more complex mnemonic structure such as a phrase or sentence (e.g., Epstein, Rock, & Zuckerman, 1960; Rohwer, 1966), (b) employing a natural language mediator such as A-Apple-Pie (e.g., Montague, Adams, & Kiess, 1966; Schwartz, 1971), or (c) embedding in a unitary visual image (Asch, Ceraso, & Heimer, 1960; Epstein, Rock, & Zuckerman, 1960; Paivio, 1969). Considering the efficacy of these mnemonic devices for paired associate learning and the typical lack of control over the learning methods used by

This research was supported by National Science Foundation Grant BMS75-03145 to the author and by Contract F 44620-73-C-0056 from the Advanced Research Projects Agency, U.S. Department of Defense, to Ray Hyman.

Requests for reprints should be sent to Wayne A. Wickelgren, Department of Psychology, University of Oregon, Eugene, Oregon 97403.

rote-rehearsal subjects, one can question whether much direct rote association of A and B ever occurs in paired-associate learning.

Third, in memory for serial lists, the importance of organization by serial-position cues and grouping structure has been repeatedly demonstrated (e.g., Bower & Winzenz, 1969; Wickelgren, 1964, 1967; Young, 1968), although there is also good evidence for the simultaneous existence of item-to-item associations in both short-term and long-term retention at a phonetic level of memory (Wickelgren, 1965, 1969, 1972). In addition, mnemonics of the form "one is a bun" and the method of loci can greatly facilitate the learning of serial lists.

In paired-associate and serial learning, more and more theorists are making use of concepts such as chunks (Miller, 1956), organization or structure (Mandler, 1968), control elements (Estes, 1972), or unitary concept representatives that have various relationships to constituent attributes (Anderson & Bower, 1973; Rumelhart, Lindsay, & Norman, 1972; Schank, 1973; Wickelgren, 1969). All models of semantic memory assume that events or propositions are represented by a hierarchical or network structure of associations between the constituents.

Geometric Increase of Connection Capacity in Associative Networks

It seems increasingly undeniable that the human capacity for associating any two ideas at a conceptual level does not result because every idea can be directly connected to every other idea. Rather, it seems necessary to assume that this capacity for associating any two ideas results from an indirect association through a connecting system of intervening nodes. As Mandler (1968) pointed out, the number of items that can be associated one to the other by means of an organizational hierarchy can demonstrate a geometric increase as the number of intermediate nodes (associative distance) increases. Although Mandler applied this principle to learning concepts in a categorical hierarchy, precisely the same principle applies as well to the association of elements via any sequence of vertical associations in a propositional network.

Hence, any of the semantic-memory systems currently envisaged also provide for a geometric increase in connection capacity with increasing associative distance (an increase in the number of intervening nodes). Thus, a semantic-memory network automatically provides for the connection of every idea with every other idea. Furthermore, it does this in a manner that represents linguistic and semantic structure.

Neural Analogues

Although all suggestions concerning the neural representation of concepts or propositions must be considered speculative at the present time, it is interesting to note that the need to provide connection-capacity between every idea in a semantic network with a limited number of links is paralleled by the need to connect every neuron to every other neuron in the cortex. It is estimated that there are approximately 2×10^9 neurons in the human cerebral cortex (Pakkenberg, 1966). At the same time, the number of synapses per neuron, while varying substantially for different neurons, is on the order of 4×10^4 for the human cortex (see Cragg, 1975). Thus, every cortical neuron cannot be directly connected to every other cortical neuron, though every cortical neuron may be indirectly connected to every other cortical neuron via one or more interneurons. In fact, just one interneuron could serve to connect every pair of neurons in the cortex, since $(4 \times 10^9) \times (4 \times 10^4) \div 2 \times 10^9$.

Internodes as Chunk Nodes

It is parsimonious to view the intervening node connecting two nodes A and B as a chunk node representing the conjunction A&B, if we make the assumption that if A (or B) is connected to A&B, then A&B is connected to A (or B). This builds in the possibility (but not the necessity) of equivalent forward and backward associations between every pair of nodes. Besides serving as a higher order chunk node, the A&B node serves as an internode connecting A to B by an *upward* association from A to A&B and a *downward* association from A&B to B.

Even if the number of nodes in memory is as large as the number of neurons in the

cortex, 2×10^9 , every pair of nodes can be directly connected to a single intervening chunk node under the assumption that the connection-capacity of each node is on the order of 4×10^4 (the number of synapses/neuron in the cortex). In combining three or more nodes into a single chunk in a connection network with 2×10^9 nodes and 4×10^4 connections/nodes, it is highly probable that three or more nodes will converge on inter-nodes (chunk nodes) in a pairwise fashion—for example, A&B and (A&B)&C. With such a network, it is highly improbable that three nodes will converge on a single node without prior convergence of a pair of nodes. Thus, network connection-capacity provides a theoretical argument supporting a hierarchical structural encoding of any multielement event, with binary branching at virtually every node in the structure. An entirely different argument in favor of binary branching was given by Anderson and Bower (1973, pp. 246–247).

Both theoretical arguments and cognitive and psycholinguistic evidence favor a network representation for the interconnection of ideas. Physiological and anatomical studies of sensory systems indicate that a considerable degree of vertical hierarchical organization occurs in the analysis of sensory stimuli. Thus, vertical association and the representation of complex sets of lower order nodes by higher order nodes appears to be a ubiquitous feature of coding in the nervous system at all levels.

PURPOSE OF NETWORK STRENGTH THEORY

If the association of A and B is not by a direct (horizontal) connection, but rather by a network of indirect (vertical) connections, can any meaning be attached to the notion of strength of association between A and B? Many researchers who assume memory to be a complex network have decided that we must give up on efforts to provide a simple abstract mathematical theory of storage and retrieval dynamics. They would either forego quantitative theorizing altogether or else make a serious attempt to "map out" the detailed structure of semantic memory and combine this structure with storage and retrieval processes in a complex computer-simulation model.

Network strength theory takes a different tack by attempting to find assumptions under which the storage and retrieval dynamics of an indirect, network-mediated association will behave like a simple direct association.

Network strength theory is similar in some of its basic objectives to a theory put forth by Giuliano (1963) as a model for word association (an abbreviated, but more accessible presentation of the model can be found in Norman, 1969). However, the specific mathematical theory is different in ways that have rather important additional consequences.

In the following section on storage dynamics, the assumptions for serial and parallel combinations of links are presented without consideration of retrieval dynamics. That is, in this section, network strength theory is described in its asymptotic form. The predictions concern the asymptotic strength of association between two nodes after two or three seconds have elapsed following the retrieval probe. I will demonstrate that under many conditions, the storage dynamics for the strength of association between two nodes has the same form as the storage dynamics for each individual component and that the decay-rate parameters for the whole are simple averages of the parameters of the component associations. Under other circumstances, the retention function for the whole bears a reasonably simple relationship to the retention functions of the components even though it is not identical in form.

In the subsequent section of the paper, the theory is generalized to accommodate retrieval dynamics and to predict the entire speed-accuracy tradeoff function for recognition or recall. The speed-accuracy tradeoff function is some measure of accuracy in recall or recognition plotted as a function of the time the subject has in which to make a response (Reed, 1973, 1976; or Wickelgren, 1975b, in press). The most novel aspect of network strength theory is that retrieval of several links in an associative chain can, under some conditions, occur essentially at the same time (in parallel). Semantic-memory theorists have debated the extent to which links diverging from a common node can be activated or searched in parallel, but it has

always been assumed to be logically necessary for two links in series to be retrieved sequentially. This is not so, and network strength theory actually permits any degree of serial versus parallel processing through a chain of associations by changes in parameter values.

The last section briefly considers some applications of the theory to explain phenomena in recall and recognition, semantic memory, speech recognition, and reading.

STORAGE DYNAMICS

Across a variety of conditions (paradigms, study times, retention intervals, delay-filling tasks), types of materials (letters, digits, nonsense materials, words, sentences, pictures, tones), and subjects (children, adults, elderly subjects, amnesic patients, intoxicated subjects), the same form of retention function may describe long-term memory for delays from 30 sec to 2 years (e.g., Wickelgren, 1974, 1975a):

$$S(t) = \lambda (1 + \beta t)^{-\psi} e^{-\pi t}, \quad (1)$$

where $S(t)$ is the strength of the trace in d' units; t is the retention interval (in sec); λ is the degree of learning (trace strength at $t = 0$); β and ψ are the parameters of the consolidation and time-decay processes, respectively (constants for a given individual under most conditions); and π is the variable parameter of the interference process representing the degree of similarity between the previously stored association and the material learned (or processed) during the retention interval. Most of the evidence supporting Equation 1 has not been published, so no pretense is intended that Equation 1 has been established as an invariant law of forgetting. We are concerned with whether this same multiplicative forgetting function *could* be obtained for a wide variety of materials independent of the structure and complexity of their network representation. There are at least two approaches to this question: (a) One could argue that the encoding of all materials involves a large network structure and could show that under certain conditions, forgetting in large networks converges toward a common form as the size of the network increases, regardless of the nature of the net-

work structure and the forgetting parameters for each link. (b) Alternatively, one could try to establish some conditions under which forgetting in networks has the same form as the basic forgetting function for an individual link. It is this latter approach that is taken here. The basic strategy is to find some possible rules for defining network strength from the strengths of the component links of the network and then to see if there are conditions in which the form of the forgetting function for the network is the simple product function given in Equation 1, assuming that Equation 1 is the form of forgetting function for individual links (but with possibly different parameter values for λ , β , ψ , and π for different links).

Series Connections of Associations in a Chain

Since it is assumed that two memory nodes are virtually always connected by a chain of two or more associative links, a rule for series combination of component strengths is fundamental to network strength theory. We shall consider three possible rules—the *minimum*, the *geometric mean*, and the *harmonic*—and point out some consequences of each.

Minimum rule. The strength (S) of a chain of n associations with strengths $\{s_1, s_2, \dots, s_n\}$ is the same as its weakest link:

$$S = \min \{s_1, s_2, \dots, s_n\}. \quad (2)$$

According to the minimum rule, the forgetting function for a chain would be the same as that of its weakest link, so the equivalence of macro and micro storage dynamics is easily obtained. The main question is just how plausible the minimum rule is. In characterizing learning in network structures, an attractively simple rule appears to be followed by most semantic-memory theories, namely, *all learning takes place at the top of some local hierarchy*. When we learn that "gliding to a stop saves gasoline" we probably do not substantially strengthen the links from features to segments, from segments to words, from words to concepts, or from concepts to familiar conceptual components as "gliding to a stop" or "saves gasoline." This large subset of our network representation of the proposition was learned long ago and, by Jost's second law, is being forgotten at a very slow

rate. What is newly learned and subject to the most rapid forgetting is the small subset of associations at the top of the hierarchy, which link the familiar conceptual compounds into a single proposition. Most learning may involve adding a rather small number of links to the top of some well-established portions of a hierarchy. If this is correct, then some newly learned weak link at the top may often dominate storage (and retrieval) dynamics and may be analogous to a rate-limiting reaction in a complex series of chemical reactions. This might make the minimum rule a good approximation even if it is not the exact rule for serial combination of strengths.

Geometric mean rule. The strength of a chain of n associations with strengths $\{s_1, s_2, \dots, s_n\}$ is the geometric mean of the component strengths:

$$S = \left[\prod_{i=1}^n s_i \right]^{1/n}. \quad (3)$$

To some, this geometric mean rule for series connection may seem counterintuitive because it asserts that the strength of a chain of associations is stronger than the strength of its weakest link. However, the geometric mean is more sensitive to the strength of the weakest link than the arithmetic mean, and it is by no means theoretically or experimentally obvious that the strength of a chain of associations should be set at the minimum component strength, independent of the strength of all other associations in the chain. This degree of dependence on the weakest link may be optimal.

Under the geometric mean rule, demonstrating that the storage dynamics for a chain can be equivalent in form to that for a single link is critically dependent upon the presumed storage dynamics for a single link. This equivalence holds when the forgetting function for a single link has a simple multiplicative form, but it is violated when the function involves additive components. For example, if Equation 1 is assumed to hold for each link in the chain, with parameters λ_i , β_i , ψ_i , and π_i that may be different for each link, then the macrodynamic forgetting function for the chain does not have the same form as the

macrodynamic function for a single link. The lack of equivalence stems from the use of the additive component $(1 + \beta_i t)$. However, if β is a constant for different links (and both β and ψ do appear to be constant across different materials), then the geometric mean rule does yield equivalent micro and macro storage dynamics (namely, Equation 1) with decay and interference rate parameters for the chain that are simple arithmetic averages of the component rate parameters and a degree-of-learning term that is the geometric mean of the component degrees of learning:

$$S(t) = \lambda (1 + \beta t)^{-\psi} e^{-\pi t},$$

where

$$\lambda = \left[\prod_{i=1}^n \lambda_i \right]^{1/n},$$

$$\psi = \frac{1}{n} \sum_{i=1}^n \psi_i, \text{ and } \pi = \frac{1}{n} \sum_{i=1}^n \pi_i. \quad (4)$$

Furthermore, for large values of t (greater than 60 sec or so, βt is sufficiently large in relation to 1 that the retention function is well fit by an equation of the form $S(t) = \lambda t^{-\psi} e^{-\pi t}$, for which the equivalence of micro and macro storage dynamics holds irrespective of the constancy of β across links.

Harmonic rule. The strength of a chain of n associations is the reciprocal of the sum of the reciprocals of the component strengths (same rule as for combining conductances in series electric circuits):

$$S = \frac{1}{\sum_{i=1}^n \frac{1}{s_i}}. \quad (5)$$

With the harmonic rule, the equivalence of micro and macro storage dynamics can still be obtained but under more restrictive conditions than for the other two rules. If the time-decay process is truly invariant for all links, which is plausible and has some (unpublished) evidence supporting it, then we get:

$$S(t) = \frac{(1 + \beta t)^{-\psi}}{\sum_{i=1}^n \frac{1}{\lambda_i e^{-\pi_i t}}}. \quad (6)$$

It is not theoretically reasonable to assume constant π_i for links at different levels in a

hierarchy, since π_i refers to susceptibility to interference and this should be greater for links at lower nodes that participate in other propositional hierarchies. However, at a macro level in the absence of highly similar interpolated learning, π is typically very small in relation to ψ , so that for retention intervals less than a day or so, the interference factor is negligible in relation to the time-decay factor. In this case, $e^{-\pi t} = 1$, and even with the harmonic rule, one would obtain equivalent micro and macro storage dynamics, namely:

$$S(t) = \lambda(1 + \beta t)^{-\psi},$$

where

$$\lambda = \frac{1}{\sum_{i=1}^n \frac{1}{\lambda_i}}. \quad (7)$$

Which of the three rules for series combination is most reasonable is an open question. The minimum rule makes the strength of a chain equal to its weakest link, the geometric mean rule makes it stronger than the weakest link, and the harmonic rule makes it weaker than the weakest link. While most semantic-memory theories have assumed that the probability of successful retrieval generally decreases with an increasing number of links in a chain (and cannot ever increase), it is worthwhile to question the necessity of that assumption. It might be functional to regard longer chains as not necessarily weaker than shorter chains. In many ways, the minimum rule seems the most plausible of the three, and for this rule, micro and macro storage dynamics are always equivalent. However, there are an infinity of possible series-combination rules and link forgetting functions, and it is premature to draw any conclusions as to which works best. The point of the present section is that it is quite possible under many conditions to develop a theory of memory storage dynamics that holds independent of the exact coding of material in memory.

Parallel Connection of Associative Chains

Whenever a subject learns two images or two propositional or single-word mediators that link components of a pair, it is plausible to imagine that there are two separate chains

of associations linking the elements of the pair. Such separate chains are assumed to be connected in parallel, by analogy to electric circuits. Another example of parallel connection arises in situations where a visual-image connection is assumed to exist simultaneously with a verbal (generally propositional) connection. Spaced learning trials appear to generate multiple traces to some extent, rather than merely incrementing the strength of a single trace (Hintzman & Block, 1971). Finally, it might be necessary to assume parallel links between nodes in memory in serial-list learning, in which it appears plausible to assume both item-to-item associations and associations involving serial position or grouping concepts.

It seems reasonable that under many conditions, it might be possible to limit the connection between two nodes to a single chain by carefully controlling the method of encoding used by each subject on each trial (independent of whether we have any exact theory of the nature of the encoding). For this reason, primary emphasis should be placed on the invariance of storage dynamics under series combination, rather than on the conditions under which this invariance might hold for parallel combination of component links. However, experimental situations certainly exist for which a limited degree of parallel connection will probably have to be assumed. Under natural learning conditions, semantic memories may involve a large number of parallel connections.

Accordingly, it is of some interest to consider a plausible rule for the combination of associations in parallel. The most plausible rule would appear to be a simple *additive combination*. Let S be the asymptotic strength of association between node A and node B, and let S_j represent the strengths of each of the parallel associative chains from A to B. Then the additive rule is:

$$S = \sum_{j=1}^m S_j. \quad (8)$$

If the additive rule for parallel chains is combined with any of the three rules for the individual chains, then one obtains a macro-dynamic forgetting function equivalent in

form to the microdynamic function only in the case where the time-decay and interference parameters are approximately equal for each chain. For the minimum rule, this means that the decay and interference parameters must be equal for the weakest links. For the geometric mean rule, it means only that the *averages* of the decay and interference parameters must be approximately equal for the different chains. For the harmonic rule, the conditions for obtaining equivalent micro- and macrodynamics for a single chain are already strong enough to yield the same equivalence for a set of parallel chains.

Note that to obtain invariance of the retention function for parallel combinations of associations, it is not necessary to impose restrictions on the magnitudes of the acquisition parameters. Also, no restrictions need be placed on the number of intervening associations forming each of the associative chains, nor are the number of links required to be equal for the different chains. Furthermore, it is quite possible to experimentally study the retention function for each of the two encodings of an association separately, and then to study the retention function for the association when both encodings are learned. This would test the validity of the additive combination rule for parallel associations.

An alternative to the additive rule for parallel connections that is analogous to the minimum rule for serial connections is the *maximum rule*: The strength of m parallel chains is the maximum of the strengths of the individual chains. As with the minimum rule for serial combinations, the maximum rule for parallel combinations always yields a form of the macro retention function for the association between nodes A and B that is identical to the form of the micro retention function of the component chains.

Compound Cues

Both the additive and the maximum combination rules for parallel connection can be naturally extended to the combination of strengths in the case of compound (multiple) cues for memory retrieval. If both nodes A and B are associated to node C by indepen-

dent chains and the additive combination rule is followed, the strength of association from the compound cue A + B to node C would be the sum of the strengths of associations from A to C and from B to C. If the maximum rule is followed, then the strength of the compound association would be the maximum of the component associations.

RETRIEVAL DYNAMICS

Studies of recognition-memory retrieval dynamics by Reed (1973, 1976), Corbett (1975), and Doshier (1976), who used the speed-accuracy tradeoff method, indicate that the time to retrieve an associative strength in a recognition-memory test is on the order of 1 to 2 sec. The speed-accuracy tradeoff function begins to rise above chance performance (at the time intercept) after about 300-600 msec, and nearly asymptotic performance is achieved anywhere from .5 sec to 1.5 sec later. Over this time, the increase in accuracy is generally negatively accelerated in a manner that is well fit by an exponential approach to limit:

$$S(T) = S(1 - e^{-\gamma(T-\delta)}), \quad (9)$$

where $\{T - \delta\} = T - \delta$ for $T > \delta$ and $\{T - \delta\} = 0$ elsewhere.

In the above equation, $S(T)$ stands for the retrieved strength of association at T seconds following the onset of the probe item or set of items; S represents the asymptotic strength of this association as retrieval time T approaches infinity; γ represents the rate of approach to the limit; and δ represents the time intercept of the function (the largest value of T at which $S(T) = 0$).

Reed prefers functions other than an exponential approach to a limit to describe his obtained speed-accuracy tradeoff function for recognition-memory data, but the exponential approach to a limit provides a reasonable approximation to these data, and it is really too early to make any definite decisions concerning the form of the speed-accuracy tradeoff function for recognition-memory retrieval. None of the conclusions that follow depend in any way upon the assumption of an exponential form for the speed-accuracy tradeoff function $S(T)$.

The only important assumption to be made concerning the form of the retrieval function is that it be the product of a time-invariant, asymptote factor and a time-varying factor representing the approach to that asymptote, with the latter factor being independent of asymptotic strength. That is, the speed-accuracy tradeoff function for any component link (association) is assumed to have the form

$$s_i(T) = s_i \cdot r(T). \quad (10)$$

In the above equation, s_i represents the time-invariant term and $r(T)$ represents the "universal" retrieval function for the increase in retrieved strength of association as a function of retrieval time T .

This is a strong assumption concerning the nature of retrieval dynamics, since it assumes that both the form and the parameters of retrieval dynamics are the same for all associations and, furthermore, that the approach to an asymptote is independent of the value of that asymptote. Nevertheless, one has a choice of an infinite variety of possible retrieval functions $r(T)$. Regardless of which universal retrieval function is chosen, one can draw a very important and novel conclusion regarding the retrieval dynamics of a chain of associations.

Chain Serial versus Parallel Processing

When a psychological process is considered to be composed of a variety of several subprocesses, the issue of serial versus parallel processing concerns whether the component subprocesses are executed sequentially (one at a time) or simultaneously (all at the same time). Whenever the component subprocesses are organized into a chain such that one end of the chain constitutes the starting point for processing and the output of each link in the chain is the only input for the next stage, everyone has previously assumed to my knowledge, that serial processing is necessarily implied. As we shall see in the present section, having a set of links (associations) organized into a chain does not necessarily imply serial processing of the separate links. In fact, the present network strength theory of retrieval dynamics actually yields a continuum of various degrees of serial versus par-

allel processing through a chain of elements, dependent upon the relative values of certain parameters. The specifics are as follows.

In general, one assumes that there is a delay of τ under the parallel processing assumption ($\tau \approx 1$ msec) between the initiation of retrieval of the first association in a chain and the initiation of retrieval of the second association in that chain. Thus, in the parallel-processing version, a speedy little activation pulse is assumed to travel through a chain of n links in $n\tau$ msec and to initiate retrieval of trace strength at each link, but the completion of this retrieval requires several hundred milliseconds. If one assumes that this nodal delay parameter τ is invariant for all nodes, then the form of the retrieval-dynamics function for a chain of n associations can be expressed in a particularly simple way as in Equations 11 through 13.

Minimum:

$$S(T) = \min\{s_i \cdot r(T - [i - 1]\tau)\} \quad (11)$$

Geometric Mean:

$$S(T) = \left[\prod_{i=1}^n s_i \right]^{1/n} \left[\prod_{i=1}^n r(T - [i - 1]\tau) \right]^{1/n} \quad (12)$$

Harmonic:

$$S(T) = \frac{1}{\sum_{i=1}^n \frac{1}{s_i \cdot r(T - [i - 1]\tau)}} \quad (13)$$

The forms of the retrieval function given in Equations 11 through 13 are not identical to those for component links, unless $n\tau$ is small in relation to the various intercept and rate parameters characteristic of the universal retrieval function $r(T)$. Making the assumption that $n\tau$ is negligible in comparison to the rate of retrieval of the associations is essentially equivalent to assuming parallel processing through a chain of associations. That is, the delay between the initiation of retrieval for each component link in the chain is small when compared to the time required for retrieval of full associative strength at each link, and therefore it can be ignored. Under such parallel-processing conditions, the

forms of the retrieval function for a chain of associations are given in Equations 14 through 16 and they are identical to the forms of the retrieval function for individual links.

$$\text{Minimum: } S(T) = \min\{s_i \cdot r(T)\} \quad (14)$$

$$\text{Geometric Mean: } S(T) = \left[\prod_{i=1}^n s_i \right]^{1/n} r(T) \quad (15)$$

$$\text{Harmonic: } S(T) = \left[\frac{1}{\sum_{i=1}^n \frac{1}{s_i}} \right] r(T) \quad (16)$$

So long as the parallel-processing assumption ($n\tau$ negligible) is valid, the theory obviously generalizes to handle parallel combinations of chains of associations, such that the total strength has the form of the product of an asymptotic strength factor multiplied by the universal retrieval function $r(T)$.

Network strength theory also makes the typical parallel-processing assumption that retrieval proceeds in parallel through all links emanating from a given node. This contrasts sharply with the serial-processing assumption made by Anderson and Bower (1973), among others.

It is a rather elegant property of the retrieval dynamics of network strength theory that it provides for the expression of any degree of serial or parallel processing by means of the value of a single parameter τ .

The equations for retrieval dynamics are simple only under the assumption of parallel processing; for example, when $n\tau$ is small in relation to the time period of speed-accuracy tradeoff for an individual association. However, the assumption of parallel processing through a chain of associations is actually quite plausible considering what we know of the properties of the nervous system. Of course, it is completely speculative to consider the nodal-delay time τ in any way analogous to synaptic-delay time. However, if one makes that leap of faith, we know immediately that τ is on the order of 1 msec, while the time period for speed-accuracy tradeoff appears to be on the order of 500 msec or more (Corbett, 1975; Doshier, 1976; Reed, 1973, 1976). Of course, these speed-accuracy tradeoff findings concern effective recognition-memory strengths that may represent complex serial

and parallel combinations of individual component strengths. Nevertheless, it would be difficult to make the speed-accuracy tradeoff functions for recognition memory consistent with a theory in which a large number of serially ordered delays τ occur prior to arriving at the final step in the chain, at which point the speed-accuracy tradeoff function rapidly reaches asymptote. A serial-process explanation of the relatively long time course of speed-accuracy tradeoff would require an assumption of very substantial random variation from trial to trial in the parameters regulating processing speed at each link in the chain.

The point of the present discussion is that the parallel-processing assumption is at least as plausible, if not more plausible, than the assumption of serial processing through a chain of associations. While chains have undoubtedly appeared to many people as an obvious example of serial processing, the present theory demonstrates that this is not necessarily the case. The underlying philosophy of the nervous system that generated network strength theory is similar to the principle of "continually available output" advanced by Norman and Bobrow (1975). This principle implies that the judged variable on which any cognitive decision is based does not suddenly go from a chance level to asymptote after the completion of the required number of serially ordered stages; rather, the judged variable rises gradually from chance to asymptote as time for processing increases. Network strength theory, augmented to include retrieval dynamics, provides an example of one precisely defined system in which the property of continually available output can be realized.

APPLICATIONS

In this section I shall briefly consider some present and future applications of network strength theory to the understanding of phenomena in four areas: recall and recognition dynamics, coding and retrieval in semantic memory, speech recognition, and reading.

Recall and Recognition Dynamics

The predictions of any particular network strength theory concerning retrieval dynamics

in recognition memory can be tested quite directly using the speed-accuracy tradeoff method described by Reed (1973, 1976). For example, in testing recognition memory for an A-B paired associate, one presents the A-B pair, and then after a variable delay (ranging from 0 to 4 sec), one presents a signal to the subject to make a *yes-no* response concerning whether the pair has been presented previously. Subjects can be trained to make their *yes-no* decision at a relatively constant latency on the order of 200 msec, largely independent of the lag between the onset of the pair and the onset of the response signal. This permits plotting of a speed-accuracy tradeoff function in which accuracy in the recognition-memory decision (measured in d' units, for example) is plotted as a function of the retrieval time allotted to the subject. In this way, the time course of retrieval in recognition memory may be assessed relatively directly.

Wickelgren and Corbett (in press) have also developed a procedure to assess retrieval dynamics in paired-associate recall using the same sort of two-choice *yes-no* response employed in recognition memory. To test recall of an A-B paired associate, one presents the A item followed after a variable lag by presentation of the correct B item or an incorrect D item. Subjects are told that after presentation of the A item, they should attempt to recall the B item; this permits them to make a rapid *yes-no* decision, following presentation of the second item, as to whether their recalled item matches the presented second item.

The purpose of the *yes-no* recall task is to assess the dynamics of retrieval of the unidirectional association from A to B without any contribution from the association in the reverse direction from B to A. Since the time intercept of the speed-accuracy tradeoff function for paired-associate recognition memory appears to be around 400 msec, and since subjects appear capable of making a *yes-no* recall response in the current procedure within a period of about 400 msec, it appears likely that pair recognition in no way contributes to any increase in accuracy as a function of increasing the delay between presentation of the A and B items (Wickelgren & Corbett, in press). Thus, there is some support for the

assumption that the *yes-no* recall procedure assesses retrieval of only the unidirectional association from A to B.

Comparison of recall and recognition dynamics is facilitated by the use of the *yes-no* response in both cases, and such data can be transformed into discriminability (d') measures of performance accuracy using the same assumptions for both recall and recognition memory. Because of the limited amount of data available, it would be premature to attempt to evaluate any specific network strength theory, but this is so only because of the absence of data that we know how to obtain, not because of any intrinsic untestability.

Coding and Retrieval in Semantic Memory

In network strength theory, nodes that are close together in terms of the number of intervening associative links are not necessarily the most strongly associated. If nodes that are connected by two intervening links have one very weak link, then the strength of the association between the nodes may be extremely weak in comparison to the strength of association between two nodes connected by five strong intervening links. Of course, all other things being equal, there is a greater chance that nodes separated by a large number of links will have a weak (low strength) link, and all three of the proposed serial rules are rather sensitive to weak links. However, two nodes that are distant in terms of the number of intervening links need not be distant in terms of associative strength. This may be a very desirable property, since it seems unreasonable to assume that nature has *always* provided close connections for ideas that we wish to have strongly associated. The opportunity to develop strongly associated nodes through a large number of intervening links, far from being a suspect property of network strength theory, is actually a desirable property.

Under the parallel-processing assumption, network strength theory also makes the somewhat surprising prediction that the rate parameters for the speed-accuracy tradeoff function in recognition-memory retrieval for a pair of items should often be invariant with nodal distance. This invariance holds as long as nodal distance (i) is small enough that the

quantity $(i-1)\tau$ remains negligible in relation to the time period for growth of the speed-accuracy tradeoff function for retrieval dynamics. Consideration of this matter tends to justify once again the assumption of parallel processing through chains of associations, since such a system permits nodes that are distantly connected to compete effectively with nearby nodes on the basis of the strength of association.

Parallel processing through a chain of associations has the further side effect that one cannot use reaction time or speed-accuracy tradeoff retrieval dynamics to determine the structure of semantic memory, as suggested by Collins and Quillian (1969), Rips, Shoben, and Smith (1973), and others. This occurs because semantic relations between nodes that are farther apart in the network structure need not require appreciably more time to verify than those that are closer together, so long as the links are equally strong in both cases. Network strength theory thus provides an explanation for why there has been repeated failure to find the expected differences in (a) verification time for subject-property propositions with different distances in a categorical hierarchy (e.g., "a canary can sing" vs. "a canary has skin") when associative strength is controlled or is counteracting semantic distance (Conrad, 1972; Rips, Shoben, & Smith, 1973; Wilkins, 1971) and (b) recognition-memory time for pairs of propositional constituents (subject, verb, object, time, location) presumed to be at different distances in the semantic-memory structure encoding the proposition (Doshier, 1976; Anderson, Note 1). From the standpoint of network strength theory, semantic memory may well have the hierarchical structure assumed by Collins and Quillian (1969), Anderson and Bower (1973), and others, but this structure will not necessarily be revealed by studies of retrieval dynamics.

Speech Recognition

Broadly speaking, speech recognition is usually considered to involve the passage of the auditory signal first through one or more levels of general-purpose auditory analyzers then through one or more levels of specific

speech analyzers (distinctive features) to the segmental (phoneme, syllable, allophone) levels, and finally to the word and concept level. These words or concepts are typically assumed to be the basic elements of nodes in semantic memory, with other (higher) nodes representing phrases and sentences.

This neat, serially ordered, hierarchical organization of the speech-perception process was rather strongly challenged by the findings of Savin and Bever (1970) and Warren (1971), as replicated and extended by Foss and Swinney (1973) and McNeill and Lindig (1973). These studies showed that under some circumstances, syllables can be recognized faster than phonemes, words can be recognized faster than syllables, and sentences can be recognized faster than words, syllables, or phonemes. An interesting consequence of network strength theory, under the parallel-processing assumption, is that a serially ordered structure of levels in the processing of stimulus input does not necessarily predict faster recognition of target elements at levels closer to the periphery. Even the extraordinary findings of McNeill and Lindig (1973), that subjects can sometimes detect a sentence target faster than a syllable target, which in turn can be detected faster than a phoneme target, in no way contradict the ordering assumption for the levels. The explanation is completely analogous to the reason why increasing the distance in terms of the number of intervening links does not necessarily have any detectable effect upon retrieval dynamics so long as the τ parameter for nodal-delay time is small in relation to the rate of approaching the limit for the retrieval function of each link.

Reading

Precisely the same considerations that apply to speech recognition also apply to reading. Analogous to the Savin and Bever (1970) reaction-time finding in speech recognition, Johnson (1975) found that subjects can identify whether a visual stimulus is or is not some target word as fast as they can determine whether a single visual letter is or is not the target letter and faster than they can determine whether a visual word does or does

not contain a target letter. This finding has been taken to challenge the assumption that the letter level of analysis precedes the word (or concept) level. If letters are analyzed before words, how can words be identified as rapidly as isolated letters, and how can letters in the context of a word be identified even more slowly than the entire word?

Once again, network strength theory provides an explanation of these findings that permits retention of the attractive assumption that graphic segment units such as letters and/or letter dyads or triads (context-sensitive graphemes) are constituents of words that precede the word level for visual stimuli. With chain parallel processing, the delay between the onset of activation of segmental (letter) units and transmission of this activation to associated word units is negligible. Since word units receive converging input from a large variety of segmental (letterlike) units, the time to reach some retrieved-strength threshold for identification of a word might easily be the same or less than the time for a segment (letter).

REFERENCE NOTE

1. Anderson, J. R. Personal communication, February 1975.

REFERENCES

- Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D.C.: Winston, 1973.
- Asch, S. E., Ceraso, J., & Heimer, W. Perceptual conditions of association. *Psychological Monographs*, 1960, 74(3, Whole No. 490).
- Bower, G. H., & Winzenz, D. Group structure, coding, and memory for digit series. *Journal of Experimental Psychology Monograph*, 1969, 80(2, Pt. 2).
- Chomsky, N. *Syntactic structures*. The Hague, Netherlands: Mouton, 1957.
- Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 240-248.
- Conrad, C. Cognitive economy in semantic memory. *Journal of Experimental Psychology*, 1972, 92, 149-154.
- Corbett, A. *Artificial categories and semantic distance in a categorization task*. Unpublished master's thesis, University of Oregon, 1975.
- Cragg, B. G. The density of synapses and neurons in normal, mentally defective and ageing human brains. *Brain*, 1975, 98, 81-90.
- Doshier, B. The retrieval of sentences from memory: A speed-accuracy study. *Cognitive Psychology*, 1976, 8, 291-310.
- Epstein, W., Rock, I., & Zuckerman, C. B. Meaning and familiarity in associative learning. *Psychological Monographs*, 1960, 74(4, Whole No. 491).
- Estes, W. K. An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972.
- Foss, D. J., & Swinney, D. A. On the psychological reality of the phoneme: Perception, identification, and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 246-257.
- Giuliano, V. E. Analog networks for word associations. *IEEE Transactions on Military Electronics*, 1963, MIL 7, 221-234.
- Hintzman, D. L., & Block, R. A. Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, 1971, 88, 297-306.
- Johnson, N. F. On the function of letters in word identification: Some data and a preliminary model. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 17-29.
- Mandler, G. Association and organization: Facts, fancies, and theories. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
- McNeill, D., & Lindig, K. The perceptual reality of phonemes, syllables, words, and sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 419-430.
- Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, 63, 81-97.
- Montague, W. E., Adams, J. A., & Kiess, H. O. Forgetting and natural language mediation. *Journal of Experimental Psychology*, 1966, 72, 829-833.
- Norman, D. A. *Memory and attention: An introduction to human information processing*. New York: Wiley, 1969.
- Norman, D. A., & Bobrow, D. G. On data-limited and resource-limited processes. *Cognitive Psychology*, 1975, 7, 44-64.
- Paivio, A. Mental imagery in associative learning and memory. *Psychological Review*, 1969, 76, 241-263.
- Pakkenberg, H. The number of nerve cells in the cerebral cortex of man. *Journal of Comparative Neurology*, 1966, 128, 17-20.
- Reed, A. V. Speed-accuracy tradeoff in recognition memory. *Science*, 1973, 181, 574-576.
- Reed, A. V. List length and the time-course of recognition in immediate memory. *Memory and 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.
- Rohwer, W. D. Constraint, syntax, and meaning in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 541-547.

- Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Savin, H. B., & Bever, T. G. The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 295-302.
- Schank, R. C. Identification of conceptualizations underlying natural language. In R. C. Schank & K. M. Colby (Eds.), *Computer models for thought and language*. San Francisco: Freeman, 1973.
- Schwartz, M. Subject-generated versus experimenter-supplied mediators in paired-associate learning. *Journal of Experimental Psychology*, 1971, 87, 389-395.
- Warren, R. M. Identification time for phonemic components of graded complexity and for spelling of speech. *Perception & Psychophysics*, 1971, 9, 345-349.
- Wickelgren, W. A. Size of rehearsal group and short-term memory. *Journal of Experimental Psychology*, 1964, 68, 413-419.
- Wickelgren, W. A. Short-term memory for repeated and non-repeated items. *Quarterly Journal of Experimental Psychology*, 1965, 17, 14-25.
- Wickelgren, W. A. Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology*, 1967, 19, 97-102.
- Wickelgren, W. A. Learned specification of concept neurons. *Bulletin of Mathematical Biophysics*, 1969, 31, 123-142.
- Wickelgren, W. A. Coding, retrieval, and dynamics of multitrace associative memory. In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
- Wickelgren, W. A. Single-trace fragility theory of memory dynamics. *Memory and Cognition*, 1974, 2, 775-780.
- Wickelgren, W. A. Alcoholic intoxication and memory storage dynamics. *Memory and Cognition*, 1975, 3, 385-389. (a)
- Wickelgren, W. A. Dynamics of retrieval. In D. Deutsch & J. A. Deutsch (Eds.), *Short-term memory*. New York: Academic Press, 1975. (b)
- Wickelgren, W. A. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, in press.
- Wickelgren, W. A., & Corbett, A. T. Associative interference and retrieval dynamics in yes-no recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory*, in press.
- Wilkins, A. J. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 382-385.
- Young, R. K. Serial learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

(Received April 12, 1976)