[all emphasis in original]

Certain kinds of symbolic representations of the past in a memory system are almost as old as life itself—because the symbols that make up the *genetic code* involve precisely such a system of representation. A symbolic representation of something is a means to assign something (a structure) that is *different from* an event, object, reaction or feeling but in which the symbol *actually comes to stand for* the experience. This kind of procedure was reinvented when symbolic communication, through speech, was evolved by humans or their immediate predecessors. Thus a word can stand for a thing, a quality, an action, or a relationship between things and events. In the genetic code, *only things* and their linear relationship in space to each other are represented (the sequence of amino acids that a genetic sequence stands for and into which it can be translated by appropriate machinery).

In an attempt to explore the possibility that the brain uses a code analogous to language or genetic codes in transmitting and particularly in storing and manipulating information, the author (Strehler, 1969) published a series of speculations in which it was proposed that the kind of code that the brain uses in such informational transactions consists of specific patterns of events in time. The events of significance were proposed to be individual nerve discharges, and the neural code was suggested to consist of specific *discharge patterns in time*—though the impulses that make up a pattern need NOT be immediately consecutive ones. Each of these patterns of discharge was suggested to stand for a different event, object, quality, etc. Although some initial findings that were highly consistent with these speculations were published in the mid–1970s (Strehler, 1976, 1977, 1983) it was only since the late 1980s that quite extensive and persuasive evidence for the existence of just the kind of coding sequences postulated was firmly established (see Lestienne and Strehler 1987–89; Strehler and Lestienne, 1986–89) as was evidence that such patterns represent memories that have a defined lifetime in the system measured (monkey occipital cortex) (Strehler and Lestienne, 1989).

The usefulness of this new kind of memory system can be understood in very simple terms. Suppose that the “key” components in some sensory input consist of three patterns called A, B, and C, where A is the pattern / //, B is the pattern ///, and C is the pattern // /. Each of these patterns would, of course, have to be generated when item A, B, and C were detected in sensory inputs. If this combination of patterns were generated *at about the same time*, it would be possible to store the fact of their occurrence in *association with* another pattern, X, consisting of the sequence of discharge / / /, in such a way that when A, B, and C were presented to the system in the future they would generate (three copies of pattern) X, and conversely that when the pattern X was presented (or re-presented) to the system it would automatically generate an output of *all three patterns* A, B, and C.

Similarly, if the patterns G, H, and I were symbols that stand for a *response* generated as a result of the input of sensory patterns A, B, and C, this combination (G ,H, I) could be stored in association with pattern Y. Patterns M + N, standing for the feelings generated by the response, could in like manner be associated with a separate pattern Z. Symbolically, A+B+C = X; G+H+I = Y and M+N = Z!

Now the simplest way that the complex sensory patterns A,B,C could be associated with complex response patterns G,H,I and with the complex of affective symbol patterns M,N is to associate the patterns X, Y, and Z (that stand for each of these separate complexes of patterns) with each other so that the input of X will automatically generate the output of Y and Z (or conversely that the input of Y or Z will elicit the output of the two other associated patterns). Then, because X can generate A,B,C, Y can generate G,H,I, and Z can generate M,N, the presentation of data that generates the output of X, Y *or* Z will automatically generate a detailed representation of the sensory input A,B,C plus G,H,I (symbols for the response pattern) plus M,N (symbols for the affect pattern).

Note that because X stands for a set of related symbols, it is actually an *abstraction* for A,B,C, and that Y and Z similarly stand as abstractions for the input symbols associated with them. X, Y and Z can, in a manner identical with the means used to store what they stand for in association with them, be stored in association with still another symbol, e.g., “I” that serves to associate X, Y and Z in memory stores. Such a system of storage is what is called a hierarchical memory system *because each (if the symbols that stand for a collection of other symbols is at a higher hierarchical level of abstraction than are the symbols for which they stand*. The structure described above is illustrated in Figure 2.

Such a scheme can only be of use to a neural system if some means exists through which an input derived from some source can cause a specific response when it is detected. For example, if the input consists of A and B, then the components of the system that were used to store A and B should specifically respond when the patterns A and B are presented at any future time. And, if A and B are detected, they should cause the generation of pattern X, which in turn should ultimately generate all of the related patterns: C, I, Y, Z, G, H, I, M, and N.

The simplest means through which a given pattern could be recognized (the essential operation in the retrieval of stored information) is for the pattern that stands for a coded (previously stored) item of information to be presented to a *delay line system* in which the *times of arrival* of separate pulses that make up a pattern are identical at *different locations* on a memory storage-decoding neuron. (This constitutes a spatial summation network—or in electronic terms is a multiple AND gate.)

To reiterate, the use of symbols to stand for simple or very complex collections of information during storage greatly *simplifies* the manipulation of information, particularly the retrieval of sets of related information. Once such symbols for essential information are retrieved, they can be used to cause the display of their contents in more literal form in some kind of appropriate display monitor. The outputs of such a monitor can then be used by the “ultimate” monitor (the locus of the self) in selecting among options derived from memory stores.

Implicit in the use of symbols are other derived capacities of such a system. Among these is the capacity to reason (perform logical operations with symbols), the capacity to generate analogies (the most singularly powerful intellectual quality of highly evolved—human—brains), the capacity to regenerate internal displays of past events in a relatively simple manner, and finally the potentiality to generate a self-representation within the system, a most useful and indispensible quality in the selection of actions by any such system.

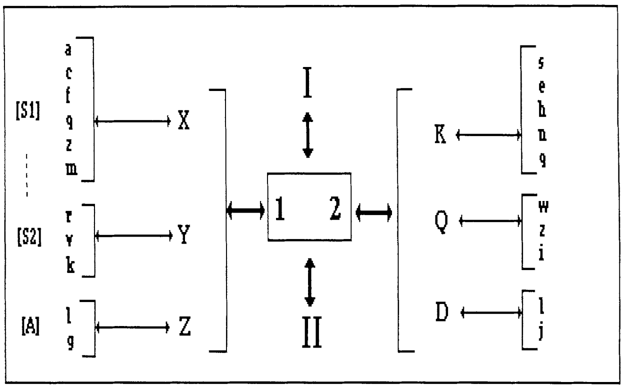
The capacity to perform logical operations (such as predicting relationships between items stored in memory on the basis of simple operations) is the easiest of these inherent properties of a symbolic memory system to describe. It is sometimes of considerable use to a living thing to “piece together” separate memories (of events that occurred at different previous times) so as *deductively* to derive knowledge that is based on relationships between things that happened at *separate times* in the past. For example, if the memory system contains a record of the fact that item A is related to item B and a separate record exists of the fact that item B is related to item C, then such a system can *deduce* that item C is related to item A. The elementary operation in deducing this *previously unexperienced or unperceived* relationship consists of inputting B and discovering that *both* A and C are related to it (because they are generated as associated outputs). Because these two symbols (A and C) are generated by the input of B, then A and C must be related to each other because they are related to a common element, B. It is, of course, not necessary to input B initially because the input of A will cause the output of B and the *re-input* of that retrieved element will cause the generation of C (and A) as an output. The important point is that although such a system has *never experienced* the relationship between A and C, the fact that such a relationship does in fact exist can be *deduced* as a result of the relationship that B has with A and C.

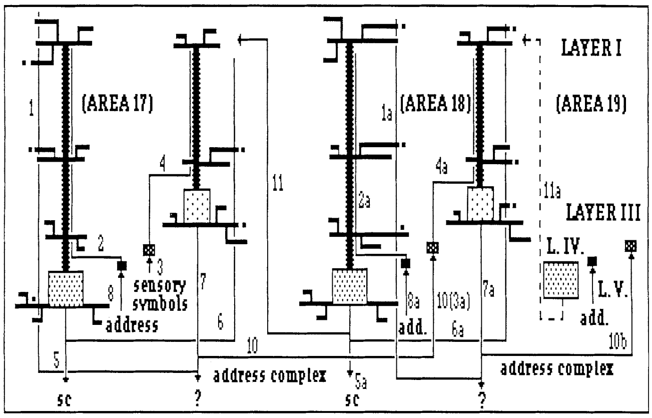
Analogy generation (such as inferring the possibility that sequential pattern Q is related to sequential pattern R) involves a somewhat more complex and less definite relationship between two complex patterns, one known (stored in memory), the other only *induced as a hypothesis* on the basis of similarities between the known pattern and an incomplete one derived from sensory information. (Such an analogy might be “The cow jumped over the moon” and “The quick brown fox jumped over the lazy dog.”) The complexity alluded to derives from the fact that complex patterns can be reduced to *abstract patterns of relationships*. In creating an analogy, the *partial* pattern in an *incompletely* understood set of abstractions is reduced to a set of abstraction symbols. This is then compared with known patterns of abstractions (or abstract relationships) that *do* exist in memory stores (or are currently being experienced) and if a correspondence is discovered between part of the abstract pattern in memory, then the newly experienced pattern is *expanded* in abstract form so that the unknown elements are filled in so as to correspond to (parallel in an abstract sense) the known pattern retrieved from memory. The completed hypothetical abstract pattern can then be translated into a literal set of relationships—a theory. Thus such an analogy machine could postulate the possibility that “The quick brown fox jumped over the moon” or the converse analogy (cow vs. dog). Bohr evidently postulated his model of the atom through analogizing heavy nuclei with the sun and the lighter electrons with the sun’s planets.

The generation of an internal display of what is stored in memory requires that a means exists through which the abstract *essence* of what is stored in memory be converted to the more literal representation required of a literal display device’s capacities. Specifically, if a set of abstract items are known to be related to a more literal, though simplified, representation of what they represent (and which they can produce), then the display that constitutes that representation can be transmitted to another display system’s monitor so as to regenerate a simplified version of what the original set of data items stood for.

This new display, derived through the manipulation of symbols, can then be made available to the decision-making center, the self, for selection of an appropriate response—the ultimate “purpose” of originally storing the information. The sets of connections needed to implement this sequence of informational transactions are evidently present in the human brain, but the specific pathways involved in this function are not yet convincingly identified.

[Figure 2] outlines some of the key features of a reciprocal, hierarchical associative memory. The upper box defines the key operations through which the symbols used are stored in association with each other in a hierarchical format. The groups of coded items in lower case letters at the left and extreme right represent combinations of specific coded symbols that are generated by various senses (S1 ... S2) and similar symbols that represent affects (A). The collections of symbols generated by the senses and affect are “bundled” together and each symbol in a group is associated with another symbol (e.g. X, Y, or Z) in such a manner that the input of a





PROPOSED CORTICAL-ANATOMICAL IMPLEMENTATION

suitable combination of the lower case (primary sensory-derived symbols, perhaps denoting features) are stored in association with each other *indirectly* through their association with the capital letters, X, Y, or Z. The symbols X, Y, Z, K, Q, and D in effect serve as “Address symbols” for the data item symbols associated with them and the large numerals “1” and “2” are addresses that are reciprocally associated with lower level addresses X, Y, and Z and with K, Q, and D, respectively.

The address symbols “1” and “2” may in turn be caused to be stored in association with some such symbol as the very large “I” (or “II") shown at the center of the diagram. Through this association the memories depicted at the right may become associated (during recall) with the memory complex on the left.

Because the system is reciprocal (which means here that the item X, for example, will cause the output of the symbols for a,e,f, ... and the input of a,e,f ... as separate items will cause the output of X), this kind of memory scheme is suited either to derive ever more “abstract” symbols when presented with cueing combinations or to generate, when operating in reverse, the particular data associated with a particular abstraction symbol.

Moreover, such a system is able to carry out the most important role of assigning meaning to a given input combination in terms of an alternative symbolic representation (as discussed in the text.) This assignment of meaning is the reason for the existence of such an evolved system, because there is little or no value to being able to recognize some sensory or recollected input *if its meaning is not also generated as associated output*.

Such an associative memory has all of the features required to deduce logical relationships *and* to carry out analogical processes when this is of value (using abstract symbols, generally) as described in the text.

The bottom box depicts a possible means through which such an associative, coded, hierarchical memory might be implemented in the cerebral cortex. Areas 17–19 of the visual cortex are schematically shown in idealized form. The lightly stippled boxes at the center represent layer 3 pyramidal cells; the more densely stippled boxes in the lower part of the diagram represent layer 5 pyramids. The small black boxes and the heavily stippled slightly larger boxes represent two kinds of input cells that lie in layer 4 of the cortex.

The dendritic stalks of the pyramidal cells have a serrated outline and several laterally projecting dendritic branches at various locations along the dendritic trunk axis, particularly at the basal and apical regions. Symbols for sensory data are presented to stellate or star pyramids in layer 4 via line 3 and then are projected via line 4, the axon of the layer 4 cell, parallel to the spines on the shafts of layer 3 pyramids, where they make (occasional?) “cartridge” synapses with the shaft (Szentagothai and Arbib, 1974). Provided that *at least 2 copies* of a given pattern are provided from the sensory system involved the cartridge synapses are temporarily facilitated, through a Hebbian mechanism, it is proposed. After such facilitation a single “scanning” impulse will, it is suggested, cause the layer 3 pyramidal cell to transmit the *mirror image in time* of the original train via line 7(1) to the dendritic branches of a nearby layer 5 pyramid. (For example, if the original train consisted of the pattern //../ then the time-inverted output of the layer 3 pyramid will be /..//. Again, provided that multiple copies of the inverted pattern are presented the synapses on the dendritic branches of the layer 5 pyramid will become potentiated (LTP). When a scanning pulse is presented at a later time the output of the layer 5 pyramid will be *identical to the original input presented to the layer 4 input cell!*

The coded address symbol input provided via line 8 to the small black stellate or star pyramid cell is processed (via lines 2 and 6) in a manner essentially identical to the processing carried out with coded sensory signals, except the roles of the two kinds of pyramids are reversed as shown. The circuitry depicted resembles a conventional electronic flip—flop circuit.

Reciprocal recall of the address upon recognition of the same sensory pattern as was stored previously, it is postulated, occurs because of the LTP of synapses previously potentiated. Moreover, the detection of a previously stored pattern will cause the read out of the associated address symbol; similarly, presentation of the address symbol will cause the read out of sensory data stored in association with that address symbol.

The pair of pyramidal cells in the right half of the diagram function in precisely the same manner as do the cells at the left with *the important difference that the input is NOT sensory data, but rather consists of a collection of addresses associated with stored sensory data symbols*. Such address symbols (e.g. the symbols X, Y, and Z in the top part of the figure) stored at synapses via line 6 are outputted Via line 10 (3a), thence via 4a, 7a and 1a.

These address symbols are stored in association with a new address symbol (e.g. “1”) via lines 2a and 6a. This process of storing combinations of addresses in association with a more abstract address symbol may be repeated a number of times similarly to the storage shown at the far right of the diagram.

The recovery of addresses at a lower abstraction level by virtue of the input of an address at a higher level is proposed to occur through a feed-back of the higher level address symbol (e.g. “1” in the upper diagram) to layer 1 where the patterns transmitted are presented, it is proposed, to the apical dendrites of layer 3 pyramids via lines 11 or 11a, which causes the layer 3 pyramids to transmit the address to the associated layer 5 pyramid for decoding (recognition). If the pattern is recognized, it causes the next lower level abstraction pattern to be emitted by layer 6 pyramids, a process putatively repeated until the original sensory data is outputted by the layer 5 cell shown at the left for output, for example to the superior colliculus via line 5.

This kind of memory circuit was first proposed in 1977 (Strehler). Many elements of it appear to be consistent with published literature on cortical connectivities. There is general consensus that sensory inputs are channeled via layer 4 cells and that these cells in turn make cartridge type synapses with the shafts of pyramids. Evidence is also available (Scheibel and Schiebel, 1970) for multitudes of recurrent collaterals of pyramidal cells, suited to make the kinds of reciprocal connections between layer 3 and layer 5 pyramidal cells posited here. Lorente de No (1949) demonstrated the kind of connections between pyramids depicted here, but whether these are rare connections or dominant ones was not reported by him. More recently, Peters and Proskauer (1980) reported one of these kinds of connections between pyramids (see also Sefton and Dreyer, 1985, p. 203). Finally, Pandya et a1. (1969, 1985) have shown that the connections between adjacent cortical areas (e.g. 17 to 18 and 18 to 19) are from upper pyramidals, including layer 3 cells to layer 4 of the next higher level. Pandya et a1. (1969, 1985) also demonstrated that layer 5 and 6 pyramids project to layer 1 of the next lower level.

References (in extract)

Lestienne, R. and Strehler, B.L. (1987a) The All-Interval Paradigm and Analysis of the Structure of Neuronal Spike Trains. Abstr. Society for Neurosciences, 13: no. 231.2, 84.

Lestienne, R., and Strehler, B.L (1987b) Time Structure and Stimulus Dependence of Precisely Replicating Patterns Present in Monkey Cortical Neuronal Spike Trains. Brain Res. 437:214–238.

Lestienne, R., and Strehler, B.L. (1988) Differences between Monkey visual cortex cells in triplet and ghost doublet informational symbols relationships. Biol. Cybern. 59:337–352.

Lorente de No, R. (1949) The structure of the cerebral cortex. In: Physiology of the Nervous System, 3rd ed. J.F. Fulton, ed. Oxford University Press, New York, 288 pp.

Pandya, D.N., and Kuypers, H.G.J.M. (1969) Cortico-cortical connections in the rhesus monkey. Brain Res., 13:13–36.

Pandya, D.N., and Yeterian, E.H.(1985) Architecture and connections of cortical association areas. In: Cerebral Cortex, Vol. 4, A. Peters and E.G. Jones, eds. New York, Plenum Press.

Peters, A., and Proskauer, C.C. (1980) Synaptic relationships between a multipolar stellate cell and a pyramidal neuron in the rat visual cortex. A combined Golgi-electron microscope study. J. Neurocytol. 8:153–183.

Scheibel, M.E., and Scheibel, A.B. (1970) Elementary processes in selected thalamic and cortical subsystems—the structural substrates. In: The Neurosciences: Second Study Program, F.O. Schmitt, ed. Rockefeller University Press, New York, pp, 443–457.

Sefton, A.J., and Dreher, B. (1985) Visual System. In: The Rat Nervous System, in G. Paxinos, ed.Academic Press, New York, pp. 169–221.

Strehler, B.L. (1969) Information handling in the nervous system: An analogy to molecular-genetic coder-decoder mechanisms. Perspect. Biol. Med. 12:584–612.

Strehler, B.L. (1977) Time, Cells and Aging, 2nd ed. Academic Press, New York, 456 pp. (see particularly pp. 324–356).

Strehler, B.L. (1983) Molecular and Systemic aspects of brain aging: Psychobiology of informational redundancy. In: Neurobiology of Aging, R. Terry and S. Gershon, eds. Raven Press, New York, pp. 281–311.

Strehler, B.L., and Lestienne, R. (1986) Evidence on Precise Time-coded symbols and memory of patterns in monkey cortical neuronal spike trains. Proc. Natl. Acad. Eci. USA, 83:981–984.

Strehler, B.L., and Lestienne, R. (1988) Presence of ghost doublets of coded neuronal patterns: Relation to synaptic memory storage. Synapse, 3:19–29.

Szentagothai, J., and Arbib, M. (1974) Conceptual models of neural organization. Neurosci. Res. Program, Bull. 12:254–402.