©1991 The Institute of Mind and Behavior, Inc.
The Journal of Mind and Behavior
Spring 1991, Volume 12, Number 2
Pages 247–262
ISSN 0271–0137

Toward a Model of Attention and Cognition Using a Parallel Distributed Processing Approach

Part 1: Background

Gregory Christ

University of Ottawa

This article reviews relevant psychology, physiology, and artificial intelligence literature, in order to present the background for a model of attention and cognition, the Sweeping Model, to be presented in a subsequent paper. Briefly described are the lack of and need for attentional mechanisms in artificial intelligence systems, followed by a description of some characteristics of human attention. Current psychological theories of attention are discussed and criticized. Physiological data about endogenous and exogenous event related potentials and attention are also presented. Finally, perception is examined with reference to its temporally discrete nature, the psychological moment.

A major task of any intelligent system, be it natural or artificial, is to select some subset of information from a universe, then analyze and integrate it into coherent structures that can result in adaptive, intelligent behaviours. The human brain is bombarded by a constant stream of sensory environmental information, only a fraction of which is useful for intelligent behaviours at any one time, with this relevant part constantly changing. If the brain, or any other system, had infinite capacity, and adequate environmental sensors, all the information in the environment could be processed fully so that no selection would have to be made. However, the brain and all other currently implementable systems have finite capacities, so if they are to operate in a complex environment, one of their fundamental tasks is to screen out the mass of stimuli that are irrelevant to intelligent behaviour. (Intelligent behaviour here is defined as behaviours that are adaptive toward reaching

---

This work was conducted first at the Cognitive Science Department at Sussex University, England, with the aid of a Commonwealth Scholarship, and was then continued at the School of Psychology at the University of Ottawa, with a Natural Sciences and Engineering Research Council Scholarship. Requests for reprints should be sent to Gregory Christ, School of Psychology, 125 University, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5.

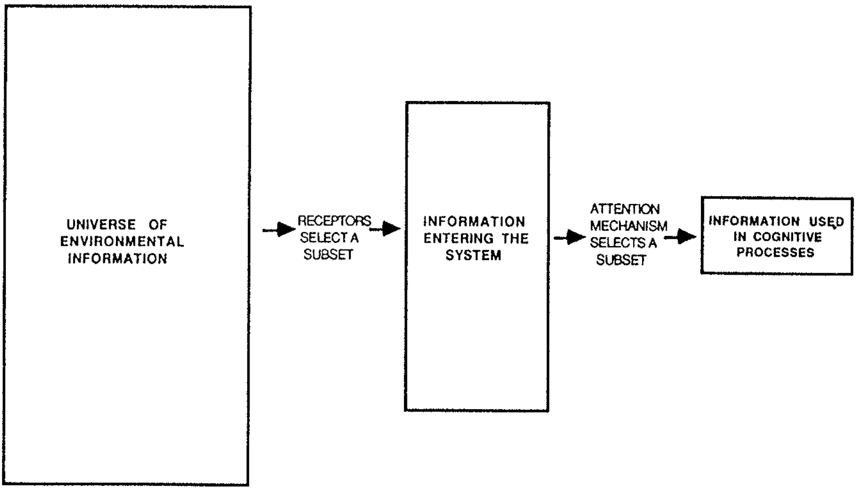

Figure 1. An illustration of how environmental information can enter a system, first by being filtered by the nature of the sensory receptors, and then by an attentional mechanism within the system.

some goal state. Goals must be postulated for any intelligent system.) Figure 1 summarizes the task of attention, and shows the two levels of the process.

Firstly, there is the infinite mass of environment information, far too much to be processed by any system's limited capacity. A small subset of this information is allowed to enter the system by the nature of the sensory receptors; that is, there are only a few sense modalities (for example, vision, audition) each of which can only sense within a certain range of stimulus energy (for example, human vision gets no information from ultaviolet light, even though it is present in the environment). Thus the sensory receptors perform the initial selection. Then, of the information entering the system, there is an attention mechanism that selects a further subset for the deeper processing required for cognitive processes; this mechanism would select which of the information would be background or focus. The model arising from the data reviewed here concerns this internal attention mechanism.

This selection of information can be viewed as one of the most fundamental cognitive processes, as it determines what information can enter the system through perception, and subsequently what can be learned, or stored in memory, and thus also what could be included in such information integrations as conscious awareness. So attention, here meaning an information selection mechanism, would seem to underlie or greatly influence all cognitive activity and behaviour. This makes some kind of attentional mechanism indispensable to any artificial intelligence (AI) systems that ever hope to approximate the richness of human cognition and behaviour. However, AI research has largely overlooked or avoided developing an adequate model of attention, and as a result may have severely restricted its progress. A major reason for this shortcoming in AI is the confusion and lack of consensus about attentional processes in human and animal psychology and physiology. However, this is no reason to ignore such a central problem to the study of intelligence and cognition when constructing AI models of these processes. There exists enough behavioural and physiological data pertinent to attentional mechanisms that a serious attempt should be made to integrate them into some theoretical framework, especially in light of new advances in AI (such as the parallel distributed processing [PDP] approach; Rumelhart and McClelland, 1986) that may give long unused observations new relevance.

Definitions

As Allport (see Claxton, 1980) points out, the term "attention" is often used synonymously with "consciousness" or "conscious awareness," which can greatly confuse what is actually under study. Here, the term "attention" will be used to refer to a selection process or mechanism that gates and regulates activity in the brain, or any other neural network. The term "conscious awareness" will refer to an overall information integration that would be synonymous with the current subjective experience of a person (a brain) or neural network. This would be determined mainly by what a human subject could verbally report being aware of, and by way of analogy to the brain, by what parts of a neural network are activated at some level. There would be degrees of conscious awareness, related to the activity levels in the neural network. Due to reliability problems of the relationship or mapping of conscious awareness with neural activity, which will be discussed later, this conscious awareness concept will be avoided except as a very general measure of brain function that must be compatible with the activity of the model to be presented.

Requirements of a Model of Attention

A model of attention adequate to explain attentional processes would have to explain and take account of several lines of perhaps loosely related evidence, such as:

(1) A wide range of behavioural phenomena. Empirical studies of behaviour would be the most important evidence for this model, because this is the only means to observe cognitive or intelligent processes.

(2) Physiological data. The model must be biologically plausible, although this is just an implementation detail and the brain is complex enough to poten-

tially support a given model in several different ways. Once an attentional model has been made conceptually clear, it could be tested by implementation using AI techniques. However, much of value in constructing such a model can probably be gained from studying the function of the brain's hardware.

(3) Philosophical concerns of how the mind is thought to work. This is a less important criterion because the current ideas could be proven inadequate, and should be changeable in light of the previous two lines of evidence.

(4) The model must agree with or explain people's subjective experience, because introspection would also be a process directed by some form of attention.

In addition to these lines of evidence, a critical factor to be considered is the ubiquitous need for sleep in biological systems, which would indicate that sleep must serve some important function. Here, sleep is regarded as another state of attention.

(5) A further requirement of any attentional or cognitive model is that it should not postulate any intelligent subsystems or homunculi. If there is ever to be an artificial intelligence, it must be achieved using intrinsically ignorant and mechanical devices; anything else might be a step in the right (or wrong) direction, but only puts off the question.

AI Models of Attention

Simple Environment

Regarding established AI models of attention, Hurlbert and Poggio (1986) make reference to the lack of and need for attention mechanisms. Little else was found in the AI literature about models of attention (for example, consult texts like Charniak and McDermott, 1985; Michalski, Carbonell, and Mitchell, 1986; Mylopoulos and Brodie, 1989). Most systems achieve the information filtering aspect of attention by having very restricted or limited input (like a simple blocks world that only deals with a few blocks on a tabletop; Winston, 1975), with all input information being fully processed. This would place the attention mechanism at the receptor level (see Figure 1), between the universe of environmental information and the information actually entering the system, with no further filtering being done. This severely limits how complex the system's environment can be for it to make intelligent responses, because either (1) much information from the environment must be missed by having very few, simple receptors, or (2) all the inputs from a very diverse environment could easily overload any existing systems, which all have some limit to their capacity. It would seem that any constructible system would have only finite capacity, which could then easily be exceeded by a complex changing environment (resulting in a combinatorial explosion of possible inputs).

Rule Based Systems (Prewired Innate Knowledge)

More recent AI systems have slightly less restricted inputs, but use heuristics or order of operations rules to determine which parts of the input receive priority to use processing resources (for example, AM and EURISKO, see Lenat and Brown, 1984; or classifier systems, see Holland, 1986). However, these heuristics or order of operations methods would have to be deliberately programmed into the system at the start and would thus correspond to some kind of innate knowledge in the human system, which is highly implausible unless these rules are extremely general, wide domain learning rules. The above mentioned systems sought to find such general rules, but were unsuccessful, even using very restricted environments. This use of much prewiring of knowledge has also been criticized for putting the "answers" into the system, then claiming that the system learned or discovered them. Schneider (1985) attempted to model attentional processes using vector transformations to represent gatings of and changes in input information, but this model too appears to have problems regarding how much knowledge must be initially programmed into the system (for example, the model uses the concept of priority learning, but does not explain how priority is determined).

Problems with Postulating Innate Knowledge

The above systems perform most of their intelligent situation appropriate behaviours as a result of the programmed-in rules and heuristics – corresponding to innate knowledge in biological systems. The problem is that these rules make the systems capable of only a restricted amount of learning in certain environments, all operating best only under circumstances forseen by the programmer. Typically, if the environment is even slightly changed from what the programmer anticipated, the system will give inappropriate behaviours – or crash entirely. This appears to happen in certain simple biological systems (for example, the sphex wasp's apparently intelligent burrow checking behaviour can be put into a continuous mechanical loop by slightly changing the environment, Woolridge, 1963). However, human beings can learn and show adaptive, intelligent types of behavior in a variety of diverse and changing environments, including ones that are highly unlikely to have been inherited as specific innate rules. This suggests that some very general, flexible, low level system for learning exists, and that the relatively highly structured rules programmed into current AI systems lack this flexibility and make these systems brittle in the face of novel, unanticipated input.

PDP's Need for Attention Mechanisms

The parallel distributed processing (PDP) approach to cognitive modelling, which is the most biologically plausible one at present, possesses no adequate attentional models (Rumelhart and McClelland, 1986, p. 114); this could be a severe hindrance to the PDP approach as the method of resource allocation and how economically allocation is accomplished could drastically change the form of the knowledge representations and capacity of neural networks. (A PDP approach is used in the model under construction, because such systems show certain properties that mirror those of the brain, such as content addressable memory, graceful degradation, spontaneous generalization, and default assignment; see chapters 1–4 of Rumelhart and McClelland, 1986).

Psychological Models of Attention

Preliminary Points

Before mentioning the main existing psychological theories of attention, some preliminary points from the research will be made, so that the theories can then be described in terms of these distinctions (for a review of the attention literature, see Johnston and Dark, 1986).

Bottom-up versus top-down processes. An important distinction, especially from an information processing perspective, is that between bottom-up or data driven processes, and top-down or internally driven processes (see Norman and Bobrow, 1975). This distinction concerns the origin of the activation or information that is controlling the attentional, or any other, process at a certain time. Bottom-up or data driven activity would involve incoming environmental information somehow controlling attention in that the physical nature of this information itself would determine which parts of that information would be filtered out or processed more deeply. Top-down or internally driven activity would involve internal structures or already stored information determining what environmental information would be rejected or selected for deeper processing. Both top-down and bottom-up processing may occur in the same model, and the interplay between the two can have important consequences concerning what information a model predicts would be processed.

Automatic versus controlled processes. Another distinction often made in attention theories is that between automatic and controlled attention (for example, Shiffrin and Schneider, 1977). Automatic attention, presumably a bottom-up process, would occur when a piece of incoming environmental information causes the attentional mechanism to give that information priority over other activity for deeper processing (that is, it attracts the focus of attention, as with the orienting reflex; Pavlov, 1927). Controlled attention would involve a limited "resource" that could be directed by "mental effort" to select specific information from environmental input, with this effort corresponding to top-down control.

Limiting resource in attention? Navon (1985) disputes the widespread notion of a limiting "resource" and suggests instead that outcome conflict or performance limitations could account for results of divided attention tasks without having to postulate a finite, shared resource. This view is supported by experiments using concurrent tasks that apparently are not limited by sharing a finite resource. For example, Allport, Antonis, and Reynolds (1972) showed that competent keyboard players could perform pieces new to them, on sight, while simultaneously repeating aloud an auditory message, with each task showing no significant difference in speed or accuracy versus when done independently. Presumably, this would be possible because the two tasks use separate mechanisms, or parts of a neural network, that would not greatly interfere with each other. Kinsbourne (1981, 1982) makes the related point that tasks are more likely to interfere the more similar they are (for example, both listening tasks), or if their processing loci are closer in cerebral space (if the areas processing each task are highly interconnected). This should be taken into account in constructing a model of attention.

Some Evidence About Attention

Shadowing experiments (divided attention). Evidence demonstrating properties of controlled attention is provided by shadowing experiments, in which a human subject must shadow (repeat aloud) one of two or more auditory messages that are presented simultaneously via different channels (for example, from different directions in a room, or more commonly with different messages in either ear using headphones). Typically, the subject can easily shadow and recall details of the shadowed message, but can recall very little or none of the unshadowed message (Cherry, 1953). This would seem to imply that attention was simply directed to one ear or the other; but it is more complicated than that, with physical characteristics of the information input being important (for example, messages into one ear could be shadowed if they differ in loudness or voice quality [Egan, Carterette, and Thwing, 1954; Spieth, Curtis, and Webster, 1954]). Cherry (1953) also found that subjects did not detect when parts of the unshadowed message were played in reverse, but they did notice if the unshadowed message's voice changed from male to female, or was replaced by a 400 Hz tone. This demonstrates that the unattended information is being processed to some degree, and that physical characteristics of this input, as opposed to the semantic content (which would require deeper processing), can attract attention in the automatic sense. Thus there appear to be at least two processes involved in attention: a controlled or top-down process that can select a certain input on the basis of physical cues and is directed by prior information concerning what those cues are; and an automatic or bottom-up process triggered by certain specific physical cues of the input that can attract the attentional mechanism to process that information further.

Main Psychological Theories of Attention

A comprehensive review of psychological theories of attention and the supporting evidence is available in chapter 7 of Martindale (1981), so to avoid repeating what has already been covered in that work, only those parts relevant to constructing a more powerful model of attention will be included in this review. The principal existing models of attention involve either early selection (in which filtering is done at or very near peripheral sensory receptors; see Broadbent, 1958), or late selection (in which filtering is done after much processing has already occurred; see Deutsch and Deutsch, 1963; or Norman, 1968). Both approaches have problems. Treisman (1969) developed a model that is a compromise between early and late selection, but it is closer to a late selection model. The primary advantage of early selection models is to account for the complete filtering out of much environmental input; and the advantage of late selection models is to explain how certain important stimuli (for example, one's own name) can attract attention over other inputs.

Galvanic skin response study against early selection models. A study casting doubt upon early selection models used conditioning of a galvanic skin response (GSR) by pairing presentations of auditory stimuli (the names of three cities) with a mild electric shock (Corteen and Wood, 1972). Thus, as would be expected from behavioural reinforcement rules, after training, when shocks were no longer given, presentation of the names of the three cities evoked a GSR in expectation of a shock. This conditioned response generalized so that presentation of other city names evoked a less intense GSR, and non-city words no reliable GSR. When words were presented in the unattended message during a shadowing task, GSR's were elicited by 37.7% of the presentations of the three previously shock-paired cities, by 22.8% of other city names, and by only 8.7% of new words. Thus, there must have been at least some semantic processing of the unattended channel in order to identify the words as city names, and early filter theories would predict gating to occur before this could happen. This finding has been replicated elsewhere (Treisman, Squire, and Green, 1974; Von Wright, Anderson, and Stenman, 1975).

Another finding of the Corteen and Wood experiment is that some processing occurs outside the focus of attention or conscious awareness, since when asked, subjects were not aware of the content in the unshadowed message or that they had been exhibiting GSR's. Thus, learning and even responding can occur outside conscious awareness, which casts severe limitations upon data gained from introspection. A model of attention would have to take these factors into account.

Points against late selection models. The above evidence, makes late selection models appear more attractive, but they also have several limitations. These limitations concern overloading the brain's, or system's, capacity or resources if all input had to be processed to a deep level before any selection could occur, which would result in the loss of much of the economy that an attentional process should provide. Also, since all input would be deeply processed, it has been pointed out that late selection models predict that shadowing should be possible from purely semantic as opposed to physical cues, but this is not the case (Broadbent, 1971).

How is the filtering done? It should be noted that these models attempt to locate the level at which filtering or selection of input occurs, but none actually describes how this filtering mechanism would work. The models point out characteristics of attention, but hide the intelligence of the system in the filter. A more complete model of attention, or intelligence, should explain how filtering is accomplished, without postulating intelligent subsystems or homunculi.

Observations. From the above evidence, it appears that there is likely a compromise between early and late selection that allows different depths of processing; that is, the selection point would not be rigidly fixed at any specific level, but would be more distributed between the peripheral sensory receptors and the higher cognitive processes, being able to use information from any level of the system to direct the process. This need for input from all levels would suggest employing a PDP appraoch to modelling human attentional processes.

Physiological Data

Physiology and PDP. At this point, it is appropriate to bring in physiological data to see what they indicate about attentional processes. A review of the neurophysiological techniques and findings with respect to learning and memory, which both involve attention, is provided by Woody (1986). That article relates various behavioural findings about learning and reinforcement principles to cellular changes in the nervous system (for more detailed description of reinforcement see Mackintosh, 1983; or Skinner, 1953). Woody also mentions that properties of synaptic weighting changes resemble "those of perceptron-like automata" (p. 459), further suggesting a PDP approach for modelling. Of course this resemblance of brain to perceptron exists because perceptrons were designed to resemble the beliefs of that time about brain functioning; Woody's comments serve to confirm that in light of newer research, this still appears to be a valid approach. Thus artificial neural networks have biological plausibility as a brain analogue, and their use is therefore appropriate for construction of a model of attention and cognition.

Exogenous and endogenous event related potentials. Concerning attentional processes, Hillyard and Kutas (1983) present an excellent summary of the physiological research. Of particular interest is the distinction between endogenous and exogenous event related potentials (ERP) as measured from the human scalp by averaging electroencephalographic (EEG) recordings taken during presentation of stimuli (usually visual or auditory) that are either to be attended to or ignored. The first 80 milliseconds (msec) of the ERP wave after a stimulus is associated with activity in the peripheral sensory pathways and is called the exogenous or stimulus bound part of the wave. The exogenous component is stimulus bound in that it varies as a function of physical characteristics of the stimulus, and is relatively insensitive to changes in information processing demands, such as whether the stimulus is to be shadowed (attended to) or not. Following this component is a larger amplitude, slower frequency wave, the endogenous component, that does change in response to information processing demands (for example, it is different for attended versus non-attended stimuli; see Hillyard and Kutas, 1983, pp. 35; 38), and indicates more activity (deeper processing?) for attended stimuli. This endogenous component, which would be associated with the controlled or effortful part of attention, has also been observed to have an approximately constant negativity per unit time over both channels in a divided attention task (Parasuraman, 1980); this suggests that there is some limited resource that must be shared to maintain attention. Just what determines the capacity of this resource and how it is allocated constitutes the main problem of controlled attention, and is further complicated by the previously cited evidence of good performance on unrelated simultaneous tasks (Allport, Antonis, and Reynolds, 1972).

Observations. The physiological data, although somewhat mixed, generally support the notion that attention can be influenced by physical stimulus characteristics (exogenous) or by previously activated internal processes (endogenous), the latter possibly drawing upon some limited resource (this resource limit may be variable rather than rigidly fixed). In addition, the endogenous attentional process appears to begin no sooner than approximately 80 msec after stimulus presentation, which is relevant to further behavioural data concerning perception.

Nature of Perception

Rapid Attentional Integrations (The Psychological Moment)

In constructing a model of attention, another factor needs to be considered – the discrete, temporal chunking nature of perception, and attentional integrations of new information.[1] This observation of the discrete nature of perception has a long history, and is summarized well in Blumenthal (1977):

> Rapid attentional integrations form immediate experience; the integration intervals vary from approximately 50 to 250 msec, with the most common observation being about 100 msec. Temporally separated events included in one integration are fused in experience to form a unitary impression; when those events are structurally different or incompatible, some may be omitted rather than fused. (p. 54)

This means that incoming perceptual information is somehow fused together or summated at input, before being allowed to enter the system in relatively discrete inputs approximately every 100 msec. All events occurring within one of these 100 msec inputs would be perceived as simultaneous. Blumenthal (1977, p. 33) discusses several lines of evidence that clarify this phenomenon, including apparent motion (for example movies), and masking effects.

In masking effects, two or more stimuli presented within the same 100 msec unit interfere with each other such that one may block out another rather than both fusing into one perception. It has been demonstrated that these effects can occur in peripheral sensory receptors (such as the retina) but also in more central brain structures. For example, a stimulus presented to one eye can mask one presented to the other eye, thus ruling out the retina as the location of that masking (Turvey, 1973). This also appears to be the case for other sense modalities. These experiments indicate that the location of summating of incoming information is at least partly within the system itself, and not limited to the receptors receiving the environmental stimulation – the receptors are capable of letting in much information that is summated at a higher level. Thus raw sensory information is present in the system, so any particular set of information can be analyzed in different ways by changes in the amount of summation, or the psychological moment (sampling rate).

Partial Report Studies

A particularly good experiment for illustrating the two-tiered nature of perception, implicating an automatic or exogenous process followed by a more flexible internally directed process, is provided by Sperling's (1960) partial report technique, also discussed by Martindale (1981), and by Lachman, Lachman, and Butterfield (1979) under the heading "iconic memory." Although many variations of the partial report technique have been conducted, the basic procedure involves tachistocopically presenting an array

---

[1] This discrete attentional chunking nature of perception is a factor apparently overlooked or dismissed by many thoeries, perhaps because of the limits of the neurological and AI technology when these theories were first formulated.

of letters (usually 3 × 3 or 4 × 4) to subjects for less than 100 msec, and then having subjects recall part or all of the array. Typically, only four or five items can be recalled, but if a specific row of letters is indicated by a high, medium, or low tone given after the stimuli have been removed, the subject can always recall the items of that row. If the tone is presented very soon after the array, it is perceived as simultaneous with it, and the target letters can still be identified even if the interval between array and tone is as large as 300 msec. However, if the tone is delayed for 500 msec, it no longer assists in recall. It is thought that this reflects the nature of iconic memory, which is related to more peripheral structures, and has a duration of about 250 msec. This iconic memory is scanned somehow by an attentional process at a maximum rate of about four to five items (for example, letters) per duration of iconic memory, or more likely, two or three different zones of the iconic field would be processed more deeply at a rate of about one every 100 msec (for example, however many letters were in each zone would be recognized). This 250 msec duration for iconic memory has been supported in replications of the partial report procedure (for example, Averbach and Coriell, 1961), as well as by different approaches testing whether visual stimuli, like circles, were perceived as simultaneous or if the icon of one had faded before the onset of another (Haber and Standing, 1969). There also appears to be an analogous echoic memory in auditory perception, which has a longer duration (Crowder and Morton, 1969; Darwin, Turvey, and Crowder, 1972; Neisser, 1967). Thus, any specific part of one of the stimulus arrays could be scanned for deeper semantic processing, but the iconic image fades before the entire array can be processed. What directs these fixations on the various iconic or echoic images of all different senses (to be collectively referred to as "sensory memory") is the attentional mechanism, which thus selects what is further processed, and a subset of this information is what will be experienced.

Nature of the Psychological Moment (PM)

It is appropriate now to briefly discuss the nature of perception and experience seen as a succession of PM's. The length of the PM is what organizes and can fundamentally change the nature of incoming sensory information and hence one's experience of the environment (for further discussion of PM's, see White, 1963). If the PM were short enough to occur a million per second, human beings could not be seen to move, falling objects would hang in the air, and one day would last an extremely long time. Conversely, if each PM lasted for a year, human beings would move too fast to be seen at all, the sun would be a strip of light continuously across the sky (with no night), small buildings would spontaneously appear, and motorways would be perceived as solid, three-dimensional objects (much as people see objects as solid rather than as subatomic particles moving around very fast). One PM contains information from the previous few PM's (that is, the prior state of activation of the network), in the form of the previous sensory activity as well as from any representations stored in memory to which this prior activation had spread. The representation that results once all the activation relaxes together contains information about changes across time, such as motion of objects – a point sometimes used to criticize distributed representations. One PM could be viewed as several successive frames of a movie superimposed on one another, with the most recent ones being the strongest; the blurred parts of the resulting image would be the parts denoting moving objects, with amount of blurring indicating the speed, and direction of motion. Concerning the neural network, the main and strongest direction of information flow is bottom-up (from sensory input from the periphery into the system), but could also occur top-down from the previously activated representations spreading in that direction (for example, during hallucinations). There is further evidence that the length of PM could be reflected in EEG frequency[2], with the commonly observed alpha-wave of about 10 Hz corresponding well with the common PM length of 100 msec (Harter, 1967).

Summary

This article has reviewed the artificial intelligence, physiology, and psychology literature, pertaining to models of attention, and cognition. The most important points to note are: (1) how attention is a basic process influencing all other cognitive activity; (2) how current models are inadequate to explain all the data; and (3) how perception and cognition occur as a sequence of temporally discrete rapid attentional integrations, or psychological moments. This information was reviewed as background for the construction of an alternate theoretical model of attention, that will involve concepts from artificial intelligence neural network models. This model, the Sweeping Model, will be presented in a subsequent paper.

References

Allport, D.A., Antonis, B., and Reynolds, P. (1972). On the division of attention: A disproof of the single-channel hypothesis. *Quarterly Journal of Experimental Psychology, 24*, 225–235.

Averbach, I., and Coriell, A.S. (1961). Short-term memory in vision. *Bell System Technical Journal, 40*, 309–328.

---

[2] EEG rhythms have been criticized as being much too general and noisy a measure to be used to infer much if anything about brain functioning. However, over billions of neurons, it is amazing that there are any relatively stable patterns at all. And since what is sought here is a very general mechanism, involving some such organized activity distributed over large portions of a neural network, EEG rhythms would be the sort of evidence that should be useful.

Blumenthal, A.L. (1977). *The process of cognition.* Englewood Cliffs, New Jersey: Prentice Hall.

Broadbent, D.E. (1958). *Perception and communication.* New York: Pergamon.

Broadbent, D.E. (1971). *Decision and stress.* New York: Academic Press.

Charniak, E., and McDermott, D. (1985). *Introduction to artificial intelligence.* Wokingham, England: Addison-Wesley.

Cherry, E.C. (1953). Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Claxton, G. (Ed.). (1980). *Cognitive psychology: New directions.* London: Routledge and Kegan Paul.

Corteen, R.S., and Wood, B. (1972). Autonomic responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology, 94,* 308–313.

Crowder, R.G., and Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics, 5,* 365–373.

Darwin, C.J., Turvey, M.T., and Crowder, R.G. (1972). An auditory analogue of the Sperling partial report procedure. *Cognitive Psychology, 3,* 255–267.

Deutsch, J.A., and Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80–90.

Egan, J.P., Carterette, E.C., and Thwing, E.J. (1954). Some factors affecting multichannel listening. *Journal of the Acoustical Society of America, 26,* 774–782.

Haber, R.N., and Standing, L.G. (1969). Direct measures of short-term visual storage. *Quarterly Journal of Experimental Psychology, 21,* 43–54.

Harter, M.R. (1967). Excitability cycles and cortical scanning: A review of two hypotheses of central intermittency in perception. *Psychological Bulletin,* 68(1), 47–58.

Hillyard, S.A., and Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34,* 33–61.

Holland, J.H. (1986). Escaping brittleness: The possibility of general-purpose learning algorithms applied to parallel rule-based systems. In R.S. Michalski, J.G. Carbonell, and T.M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach, volume II* (pp. 593–623). Los Altos, California: Morgan Kaufmann.

Hurlbert, A., and Poggio, T. (1986). Do computers need attention? *Nature, 321,* 651–652.

Johnston, W.A., and Dark, V.J. (1986). Selective attention. *Annual Review of Psychology, 37,* 43–75.

Kinsbourne, M. (1981). Single channel theory. In D.H. Holding (Ed.), *Human skills* (pp. 65–89). Chichester, England: Wiley.

Kinsbourne, M. (1982). Hemispheric specialization and the growth of human understanding. *American Psychologist, 37,* 411–420.

Lachman, R., Lachman, J.L., and Butterfield, E.C. (1979). *Cognitive psychology and information processing: An introduction.* Hillsdale, New Jersey: Lawrence Erlbaum.

Lenat, D.B., and Brown, J.S. (1984). Why AM and EURISKO appear to work. *Artificial Intelligence, 23,* 269–294.

Mackintosh, N.J. (1983). *Conditioning and associative learning.* Oxford: Clarendon Press.

Martindale, C. (1981). *Cognition and consciousness.* Homewood, Illinois: Dorsey Press.

Michalski, R.S., Carbonell, J.G., and Mitchell, T.M. (Eds.). (1986). *Machine learning: An artificial intelligence approach, volume II.* Los Altos, California: Morgan Kaufmann.

Mylopoulos, J., and Brodie, M. (Eds.). (1989). *Readings in artificial intelligence and databases.* Los Altos, California: Morgan Kaufmann.

Navon, D. (1985). Attention division or attention sharing? In M.I. Posner and O.S. Marin (Eds.), *Attention and performance XI: Proceedings of the eleventh international symposium on attention and performance* (pp. 133–146). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Norman, D.A. (1968). Toward a theory of memory and attention. *Psychological Review, 75,* 522–536.

Norman, D.A., and Bobrow, D.B. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44–64.

Parasuraman, R. (1980). Effects of information processing demands on slow negative shift latencies and N100 amplitudes in selective and divided attention. *Biological Psychology, 11,* 217–233.

Pavlov, I. (1927). *Conditional reflexes.* London: Oxford University Press.

Rumelhart, D.E., and McClelland, J.L. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge, Massachusetts: MIT Press.

Schneider, W. (1985). Toward a model of attention and the development of automatic processing. In M.I. Posner and O.S. Marin (Eds.), *Attention and performance XI: Proceedings of the eleventh international symposium on attention and performance* (pp. 475–492). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Shiffrin, R.M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Skinner, B.F. (1953). *Science and human behavior.* Toronto: Macmillan.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 7,* (whole no. 11).

Speith, W., Curtis, J.F., and Webster, J.C. (1954). Responding to one of two simultaneous messages. *Journal of the Acoustical Society of America, 26,* 391–396.

Treisman, A.M. (1969). Strategies and models of selective attention. *Psychological Review, 76,* 282–299.

Treisman, A., Squire, R., and Green, J. (1974). Semantic processing in dichotic listening: A replication. *Memory and Cognition, 2,* 641–646.

Turvey, M.T. (1973). On peripheral and central processing in vision: Inferences from information processing analysis of masking with patterned stimuli. *Psychological Review, 80,* 1–52.

Von Wright, J.M., Anderson, K., and Stenman, U. (1975). Generalization of conditioned GSR's in dichotic listening. In P. Rabbitt and S. Dornic (Eds.), *Attention and performance V: Proceedings of the fifth international symposium on attention and performance* (pp. 194–204). New York: Academic Press.

White, C.T. (1963). Temporal numerosity and the psychological unit of duration. *Psychological Monographs: General and Applied, 77*(12), (whole no. 575).

Winston, P.H. (1975). Learning structural descriptions from examples. In P.H. Winston (Ed.), *The psychology of computer vision* (pp. 157–209). New York: McGraw-Hill.

Woody, C.D. (1986). Understanding the cellular basis of memory and learning. *Annual Review of Psychology, 37,* 433–493.

Woolridge, D. (1963). *The machinery of the brain.* New York: McGraw-Hill.