Preliminary Discussion of the Design of a Large-Scale General-Purpose Neurocomputer

Richard B. Wells
MRC Institute
The University of Idaho
Nov. 14, 2003

Abstract: This paper is a preliminary discussion of a new paradigm for the general architectural, information representation, operational and design strategy for a biologically-inspired, general purpose neurocomputer based on pulse-coded neural network methods. It presents both hardware and foundational "psychological" machine structure, and discusses key issues raised by this new paradigm. The model is based on experimental findings from both neuroscience and from developmental psychology, and proposes that the key differences between a general purpose neurocomputer and a general purpose digital computer lie in the operational characteristics of the machine that come from attempting to mimic the processes discovered by these sciences. Like most discussions of computer architecture, the model presented here is generally qualitative rather than quantitative. It does however provide a general framework for subsequent quantitative and mathematical research.

I. Introduction

A. Technology. The artificial neural network has been a subject of widespread engineering interest from the time it was first shown that networks of relatively simple computing elements were capable of expressing any finite logical expression [1]. In the hands of von Neumann and others, the McCulloch-Pitts neuron model led directly to the design of the stored-program digital computer [2]-[3], which stands as the most successful and widespread achievement of biologically-inspired system design to date. The enabling factors that led to the success of the digital computer, despite its early high costs and large space and power requirements, were the applicability of the computer hardware design to a wide range of problems and the relative ease with which it could be used by persons who were not trained in the engineering aspects of its design. The digital computer is a truly protean invention: Given enough time and enough memory, any computer can solve any mathematical problem for which human beings can provide an algorithm. Its relative ease of use and the domain of problems to which it can be applied were primary factors in its commercial success over the earlier analog computers, such as those used by the U.S. Navy for target tracking and fire control during the Second World War. Its biomimetic underpinning was in part responsible for its characterization by the press and general public in the 1950s and early 1960s as an "electronic brain."

After von Neumann's early death in 1957, analog neuron models and self-adjusting learning algorithms were developed independently by Rosenblatt [4] and by Widrow and Hoff [5]. These models, known as the perceptron and the ADALINE, became the basis for the "generic connectionist neural network model" (GCNN) that today dominates the technical literature on neural networks and is the neural network model most often used in cognitive psychology research [6] and medium-scale computational neuroscience [7]. The popularity of the GCNN is due in part to the relative ease with which GCNNs can be programmed to run on the digital computer, and the vast majority of all GCNNs reported in various applications have been software-based networks. These have the advantages of low-cost implementation (they can be run on virtually any modern personal computer) and accessibility by a broad population of users whose training lies in areas other than electronics engineering or computer science. On the other hand, software implementations of neural networks are very slow because all potential advantages of the parallelism inherent in a neural network are lost in software implementations because the digital computer executes its computations serially. Consequently, very large software-based GCNNs tend to be impractical because of the enormous amount of computation time they require. This has led researchers in computational neuroscience to adopt a different type of neural network model, the population density model [8]-[9], for large-scale modeling of neuronal systems. Using this approach, networks with up to 20,000 neuron-equivalents have been constructed in software. The price paid here is loss of any direct relationship to cell-level neurobiology.

To date, few large-scale hardware neural networks have been reported, although a few small-scale networks and some networks employing a hybrid approach using a few multiplexed hardware neurons plus software (firmware) running on digital signal processors have been reported and, in some cases, are commercially available. These are at present limited to a few tens of neurons implemented on one integrated circuit chip, although pulse-coded neural network circuits involving a few hundred neurons are almost certainly going to become available over the next few years. The present small size of these networks limits their applicability to a narrow range of possible problems. Larger scale problems can in principle be worked by using multi-chip architectures. However, there are some very important and difficult packaging and other technical issues that must first be solved before such neurocomputers can become practically viable [10]. Chief among these are the problems of inter-chip connection, synaptic weight storage, and on-chip learning algorithms. This will be discussed later in this paper. The paradigm presented in this paper is a new approach to solution of these problems based on a biomimic hardware architecture and a functional architecture for knowledge representation and information processing derived from extensive research findings in experimental psychology.

Small-scale neural networks, particularly those networks implemented in software or firmware, have been successfully employed in a limited number of commercial applications. (The research literature, of course, has reported a substantially larger number of cases of neural network applications; however, these are in large part not commercialized applications but, rather, somewhat standard and well-known problems used to evaluate neural network techniques – e.g. the parity-N problem, the spiral problem, the traveling salesman problem, the retina problem, and various mapping or control problems). Probably the largest classes of engineering neural network applications have been in pattern and/or character recognition, voice and speech recognition, image processing, image compression, target recognition, traffic control, electrical power networks, industrial electronics, control systems applications, computer-aided design [11]-[12] and robotics. It is presently thought that hardware neural networks also have great potential for application to neural prosthetics, although at present the enabling technology for this still has a long way to go [13]. Neural networks have also been shown to be applicable to estimating solutions for partial differential equations and other purely mathematical problems [14].

On the other hand, so long as the problems that can be practically addressed by neural networks are limited in scope by either calculation times (for software networks) or by hardware limitations (for hardware and/or hybrid neural network chips), widespread commercial application of neural networks faces brutal competition from other well-developed engineering techniques. For one example, it is not presently clear that there is any advantage in employing small neural networks in the majority of commonly-encountered control systems applications. Here neural networks face stiff competition from both classical and modern control systems techniques as well as from fuzzy-control approaches, and no dispassionate evaluation of the present state of the art could conclude otherwise than that neural networks are presently something of a sideshow in this arena. In part this is due to the natural preference of control system designers to stick with the methods and practices they know best, and this is more of a psychological and educational issue than a technology issue. But in part the limited success of neural networks in the field of control systems is due to the fact that the types of extremely difficult potential control systems applications where neural networks have the best chance to beat standard techniques are presently avoided by the control systems community because of: 1) lack of tractable theory for attacking these problems; and 2) lack of cost-effective hardware for implementing solutions to these problems (were the solutions known). With regard to (2), there is a role that neural networks could play because neural networks do excel in coming up with solutions to ill-posed and difficult problems for which human designers find it difficult or impossible to specify a mathematical or algorithmic solution. The role they could play even at present, were a larger community of engineers versed in neural network techniques, is that of computer-aided design tool. Provided that the solution, once known, does not require real-time adaptation, neural network techniques could be used to find the problem solution, after which this solution could be transformed into a more cost-effective non-neural network solution. This, for example, is the case for the laser modulation template optimizing neural network reported in [11]-[12]. Generally speaking, *any* neural network solution that converges to fixed weights with quantized inputs and outputs, and does not require subsequent learning and adaptation capabilities, can always be mapped to an alternate implementation because the stable, fixed-weight neural network in this case merely implements a mapping function from input to output, and such functions can be implemented using standard, highly cost-effective digital techniques, e.g. programmable logic array (PLA) look-up tables or field-programmable gate arrays (FPGAs).

Possibly the greatest commercial potential for neural networks lies with problems that are presently too complex and too poorly understood for standard engineering theory and practices to handle, and which operate in environments where continual adaptation and unsupervised concept learning "in the field" or "on the job" is required. An example of this is provided by the Intelligent Traffic Control initiative currently sponsored by National Science Foundation and the Federal Highways Administration. Another example is provided by the problem of recognizing faces-in-the-crowd for airport security and anti-terrorism applications. A third example is the problem of designing mobile robots capable of locomotion, stealth, and stalking skills comparable to that of living organisms. The spinal sensorimotor control system displays characteristics reminiscent of variable-structure switching control systems [15], a specialty field within control systems theory most often applied in highly nonlinear or parametrically-uncertain control systems [16]. Small, light-weight, low-power neural networks capable of real-time operation, are in principle very well suited for this application. These commercial interests are, of course, in addition to the potential for contribution to our scientific understanding of human cognition and psychology, and to neuroscience and its applications in the health fields.

It can also be reasonably expected that commercial availability of large-scale, cost-effective neurocomputer hardware will spur many additional applications not presently contemplated, in much the same way that the availability of affordable computers led to an explosion in computer applications in all walks of modern life. Little or none of this promise is likely to materialize, however, so long as neural network technology is limited to the small-scale applications of today. Partly this is going to be due to a lack of imagination. Many of the early computer pioneers did not think that more than a few dozen computers would ever be needed; they never foresaw how technology would some day drive down the cost and size of computers, and thereby give birth to the multitude of then-unimagined applications we see today. The history of the computer presents us with a vivid lesson: **It is not necessary to foresee all the benefits and applications that a technology will inspire in order to justify investment in the development of that technology**. Early applications can and do play an important role in the evolution of the definition of a new technology, but no important technology remains limited to those applications. In the case of neural network technology, even a casual review of the engineering literature reveals that a certain lack of vision currently limits the scope and accomplishments of neural network research. The vast majority of the literature is given over to relatively minor variations on a theme, with the same or very similar types of problems appearing again and again. Indeed, the bulk of the engineering literature is given over to small-scale problems of the sort that some critics of neural networks dismiss as "toy problems."[1] One of the most important issues in neural networks is the scaling problem in training large-scale networks [84], but it is still very difficult for most researchers to get funding to explore anything but work aimed at a specific, small-scale application with an easily defensible promise of successful payoff. One leading expert in the field recently commented to this author that application-oriented research in neural networks is favored over basic research because little funding is available for anything except applications-oriented research [17]. In the opinion of this author, it is no coincidence that the most important *paradigmatic* innovations in neural network theory over the past thirty years have come from outside the field of engineering (most notably from cognitive neuroscience).

B. The Paradigm Issue. From the very beginning of the field, the Holy Grail of neural network theory has been to unlock the secrets of human understanding, cognition, and reasoning, and then to make use of this knowledge to develop engineered systems that could truly be called "cognitive" or "intelligent." In the atmosphere of logical positivism and the influence of behavioral psychology that prevailed in the 1950s, the era of the "electronic brain," it was first thought that the achievement of these goals would soon be within reach. Nearly half a century later, the first of these goals is yet to be attained and the second of these goals seems to this author to be almost as far out of our reach as at the beginning of the field. In the meantime, both logical positivism and behaviorism have rightly fallen out of favor, but the early neural network paradigms based on them have largely not changed within the broader engineering community[2]. The place these views once occupied has to some degree been filled by modern cognitive psychology and by the tremendous advances in our knowledge of neuroscience, but the larger part of the engineering world is abysmally uninformed of the most important recent developments in either field, especially in regard to the latter. Judging from the majority of the engineering literature on neural networks, many of the researchers working on neural network applications seem to have never left 1986.

To be sure, there have been many systems developed over the years that have been characterized as "smart" or even "intelligent" by their producers; but compared to the vision of the 1950s, these systems scarcely deserve that title in the opinion of this author. Large-scale neurocomputing is in one respect an optimization problem involving a very large number of input and output variables (e.g. the problem of dynamically solving the Bellman equation [85]). Only a relative few researchers elect to take on this problem. There is a paradigm problem that hampers progress in this arena, and while there are a number of different views on how such psychological objects as "concepts" and "thinking" could be pragmatically defined, e.g. [18]-[19], a dispassionate review of the literature

[1] It should be noted that whenever a large and complicated problem is broken down into smaller "blocks", these "blocks" (or their sub-blocks) eventually reach the "toy problem" level of complexity. Successful engineering is based on this sort of problem decomposition, and so the criticism of neural network research on this basis is more or less ill-founded. Criticism of "toy problem mentality" in research, however, is legitimate when such work lacks a statement of larger-scale context.

[2] This criticism does not strictly apply to the fields of cognitive and computational neuroscience, but neither of these fields is particularly concerned with neurocomputer engineering.

can only lead one to conclude that the field still lacks a common point of view even at the definitional level. The mainstream engineering neural network literature rarely contains much of any discussion of this issue, being mainly limited to narrow and *ad hoc* technical "definitions" that might permit a system to be called "intelligent" but which utterly fails to sway informed opinion in the broader scientific and technical communities. Sometimes a few rare papers make some passing reference to "results from psychology" [19], but even then the connection made is vague and lacks convincing argument to show that the model employed is anything more than a caricature of psychological findings. Jean Piaget, regarded by many as the greatest psychologist of the 20th century, once commented, "The unfortunate thing for psychology is that everybody thinks of himself as a psychologist" [20], and perhaps this is part of the problem.

In this paper is proposed a different conceptual framework for looking at this paradigm problem. The paradigm for neurocomputer operation proposed here is based in part on a more up-to-date neuroscience view of the systems-level structure and operation of the central nervous system, in part on theory that has emerged from developmental psychology, and in part on an epistemological model of human reasoning that can trace its roots back to the work of Kant [21]. Brevity demands that these constituents of the approach championed here be treated merely summarily; each of these constituent parts rightfully requires an extensive discussion of its own in order to give a full explanation. But these treatments should be done elsewhere than here, and in the first two cases have been (cf. the various citations in this paper). As for the third part and its synthesis with the first two, this part of the paradigm is the outcome of this author's own research, the notes and findings of which are not yet in publishable form [71]. Without prejudice to the usefulness that software-based modeling has for the architectonic presented here, the discussion will be centered on considerations tied to hardware implementations of this neurocomputer paradigm for the reasons previously discussed. This architecture is based on the idea of a recurrent "network of networks" architecture employing pulse-coded, time-locked, multiregional networks first proposed by Damasio [45], [49].

II. Parallels and Differences Between Digital Computers and Neurocomputers

The functions of a digital computer and a neurocomputer are in one sense the same. Both are computing devices. Both are called upon to process given information and produce particular outcomes from that information. The differences between them emerge from the types of computing problems they are called upon to solve and the level of "self-reliance" expected of the system in obtaining solutions. As we make our comparisons between them, however, it is important to realize that when we speak of the differences between them in regard to the problems they are called upon to solve, the fundamental differences have less to do with the mathematics or logic of the problems than with the operational characteristics of the problem, the amount of prior problem analysis and reduction that has taken place before the problems are presented to the machine, and with the time constraints placed upon how fast the solution must be presented if it is to be of use in the application toward which the problem-solving task is directed. The general purpose digital computer is truly general; given enough time and enough memory a general purpose digital computer can solve any problem for which an algorithm can be stated. It follows from this that if time-to-solution and memory storage requirements are not limiting factors in the problem environment, and if suitable algorithms for the problem solution are known, then the digital computer can be made to do anything that a neurocomputer can do in the context of these types of particular problems. This is self-evident when we consider that the great majority of neurocomputer operations that have been carried out to date have been executed in software running on general purpose digital computer hardware.

There has always been two primary motives for investigating neurocomputing. The first is a purely scientific motive, namely the understanding of the human brain and the phenomenon of mind. The presupposition is that if we can understand the principles of neurocomputing using artificial devices, then this will help us to understand brain and mind, thus to better understand ourselves. The scientific pursuit of this understanding is a goal of computational neuroscience, and it follows from this that neurocomputer architectures useful for this purpose should be based as closely as possible upon our understanding of the neurobiology of the central nervous system and not merely on qualitative and putative models from cognitive psychology. This scientific goal does not constrain the type of problem to be solved but, rather, the method by which the solution is to be obtained. When this general goal is divided between neurological interests and purely psychological interests, and attention is directed at the latter, the hardware platform upon which the information processing is performed is less important than the algorithmic concepts of the problem solution. This is the case for investigations into symbolic processing and traditional artificial intelligence research, and this research is invariably carried out using the general purpose digital computer as the hardware platform. Whatever the benefits of such research may be, the separation of algorithms from hardware in this case makes the problem solution rather far removed from neuroscience proper because the knowledge obtained from this sort of artificial intelligence research does not directly implicate brain function.

The second motive for neurocomputer research is purely technical. It has long been supposed that in whatever sense the word "computation" can be applied to the central nervous system, this computation is both distributed and parallel. Neuroscience findings support this supposition. The computer science interest in neurocomputing, like the interest in its potential for commercial applications, is predicated on the supposition that parallel distributed processing should lead to faster problem solution, i.e. to increased computer performance. Along with this interest there is a closely-related pragmatic interest. Human beings, as well as other species of animals, have demonstrated a remarkable ability to deal with unstructured and highly variable environments to a degree far greater than any machines have yet demonstrated. The ability to deal with such unstructured and variable environments, and to appropriately respond to changing conditions, is one definition of the word "intelligence," and the quest to understand and develop "intelligent" machines is aimed at the cost-effective achievement of this sort of operational character in man-made devices. The speed with which decisions are reached and actions are initiated in response to emerging circumstances in an unstructured and variable environment is a dominant performance factor in the design considerations of an "intelligent" machine. In some problem-solving circumstances – e.g. figuring out how to cross the street without becoming a hood ornament – human beings demonstrate speed of information processing, decision making, and action-response that far exceeds the capacity of man-made devices reported so far. It is supposed that machines with parallel distributed processing capability are better suited to attaining the objectives of this sort of problem solving than are machines that use serial processing.

Standing midway between the two distinct interests just described, and linking them, is a relatively new arena of biomedical engineering endeavors. This is the application of neurocomputing ideas to the design of practical prosthetic devices to replace lost or damaged organic, and, particularly, neural, functions. The quest for a functional and practical neuroprosthetic eye is one example of this sort of application. Artificial limbs and prosthetic spinal cords are two more examples. Another, and at present wholly speculative, prosthetic application is the replacement of lost brain functions resulting from injury or disease. In applications such as these, the neuroscience interest and the engineering interest converge because the prosthetic device must operate in conjunction with the biological organization of the human patient and do so in a manner that does not require the "re-design" of the human patient or adversely impact the quality of life in a manner worse than the original injury or disease has already done.

Finally, the majority of artificial neural networks reported in the engineering literature have been applied to the solution of signal processing, image processing, and control systems problems for which the nature of the application tends to produce mathematically ill-posed or highly data-dependent problems. Examples of this include classification problems, image segmentation, non-linear control, image noise filtering, etc. Here the typical problem solution is usually difficult to analyze in closed form but can be achieved through a "learning" process by training a small neural network on a representative set of examples where the desired input-output relationship can be specified. The success of these networks depends upon the ability of the neural network to generalize, from the training examples it has been shown, such that it usually produces correct (desired) results from input data it has not previously been shown. "Learning" processes, as they are commonly called, in such dedicated artificial neural networks generally do not carry any cognitive or other psychological implication, and we have neither any epistemological nor ontological justification for regarding networks of these kinds as "intelligent." In this paper they are regarded as merely signal processing or controller mechanisms, and as such they are merely special purpose processors and not general purpose neurocomputers. We will call them "neurocalculators."

In short, the really fundamental differences between digital computers and neurocomputers are not so much in the logical or mathematical functions they are capable of implementing as it is in the application and intent of the problem solutions to which they are to be employed. This key distinction should always be kept in mind during the step-by-step problem reductions made during research and development endeavors in neurocomputing. It is perhaps true, judging from the majority of neural network research reported to date, that this distinction is in some sense of lesser importance for small-scale problems. But it is, or should be, a primary factor in the considerations involving large-scale-problem neurocomputing devices at a level of general applicability comparable to that of the digital computer. It is from this application-oriented perspective that we need to view and evaluate neurocomputer architectures based on any biomimetic/psychological scheme.

Having said this, let us consider the factors that digital computers and neurocomputers must have in common, and what differences there are in the operational details of the two. We compare digital computer architecture with a neurocomputer architecture under the paradigm of time-locked, multiregional neural networks. At the highest level of digital computer organization there are five functions that are separable into distinct functional "blocks." These are: control function, logical-arithmetic function, memory function, input function, and output function.

A. Input and Output Functions. The input and output functions serve to connect the computer to its environment, and a parallel between these two functions and the sensory and motor functions of biological neural networks is obvious. A neurocomputer must also interface with its operating environment, and so it is clear that it requires both the input function (which we will call its sensibility) and the output function (which we will call its motor function). Obviously the specific details of the input function and the output function depend in both cases on the nature of the application environment of the machine. In the case of the digital computer, the bulk of the information "pre-processing" and "post-processing" carried out on input and output information is parceled out to the computer's "peripheral devices" and consequently are functions typically distinguished from the "central processing unit" (CPU) of the machine. Thus, although many digital computers do have specific I/O instructions, some machines have been successfully built using the idea of "memory-mapped I/O" in which input and output devices are treated in the same manner as the machine's main memory function. In this paper the broad view of the digital computer is adopted, i.e. "digital computer" is a phrase used here to denote the entire digital computer *system*. Any "boundary line" drawn between CPU function and input-output information processing functions must be seen as fundamentally a merely nominal engineering convention. For the computer system to work, consideration must always be given to all the information pre- and post-processing requirements establishing the system's connection to its environment.

When our interest is given to "intelligent" or "cognitive" systems, the relationship between input and output functions and the so-called "reasoning ability" of the machine are linked much more closely than is the case for the digital computer architecture. Woods [55] has argued that an "intelligent agent" must be capable of constructing for itself a "world model" determined by a built-in set of "beliefs, goals, and objectives" in interaction with its "perceptions" of the world in which it is embedded. It is further tasked with the job of keeping its world model sufficiently consistent with its "external world" so that it is able to achieve its goals. The actions taken by the intelligent agent affect the external world, and so the time course of the information processing it carries out is influenced by "perceptual" feedback of the changes that are taking place in the "external" world. Woods proposes the "reasoning loop" shown in figure 1 as a high-level description of this "world modeling" task.

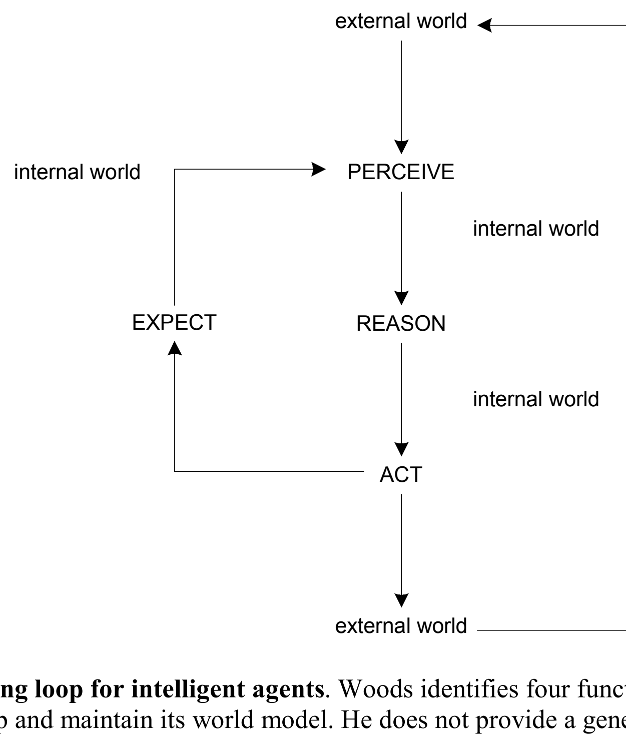


Figure 1: Woods' reasoning loop for intelligent agents. Woods identifies four functions required in order for an intelligent agent to develop and maintain its world model. He does not provide a general operational definition of these four "mental" processes (PERCEIVE, REASON, ACT, EXPECT). However, the EXPECT operation is tasked with providing a representation of anticipations of the outcomes of the agent's actions, and therefore it links the ACT function and the PERCEIVE function.

In Woods' model the PERCEIVE and ACT functions are the input and output functions, respectively, of the intelligent agent. His usage of psychological terminology here is to be taken as implying that the input function is not tasked merely with the bringing in of some data representations but additionally with forming these representations into representations of objects (entities and events). Likewise, the ACT function expresses not merely the delivery of output data representations but also implicates a sort of "psychological causality" capacity of the machine, i.e. it implicates that the machine can distinguish between changes in the external world caused by its own actions vs. those changes that occur in the external world through agencies not directly under the control of the machine. Artificial perception and machine action are linked by the ability to represent "expectations" or anticipations of changes that are to result in the external world from the machine's actions. EXPECT processes constitute a kind of "internal machine perception" distinct from the machine's "sensory modalities" that process input data from the external world. The PERCEIVE and EXPECT capacities he calls for correspond in a general way with what Kant called the capacities of receptivity and spontaneity, respectively, in *Critique of Pure Reason* and his other works [21]. It is the incorporation of these "psychological" functions that chiefly distinguishes the input and output functions of a neurocomputer from those of a digital computer.

B. Memory Function. It was recognized very early in the development of the digital computer that a means had to be provided for the storage and recall of information [3]. Furthermore, it was recognized that both logical-numerical data and data representing operational machine instructions had to be stored and recalled. The distinction between "codes" (machine instructions) and logical-arithmetic data representations was an early contribution made by von Neumann et al. This information storage and recall function was given the name "computer memory," although von Neumann was very quick to point out that "computer memory" and "brain memory" functions were quite different in nature, and that "memory" in the brain was likely to involve changes in the stimulation conditions that activate various neurons [2]. In contrast, computer memory refers to the state of binary storage elements, i.e. a flipflop "holds" either a "1" or a "0". Thus, what is "remembered" in computer memory is a datum rather than an activation condition. Furthermore, the term "memory" as it is used in reference to computers is in fact a metaphor because the term "computer memory" carries none of the psychological connotations involved in speaking of the phenomenon of memory as it is found in human beings.

Artificial neural networks, on the other hand, have from the beginning been designed to operate more along the lines of "memory" regarded as change of stimulation conditions of the neurons. This is what is meant when a neural network theorist says that "the information is stored in the weights." It also underlies the distinction made in a number of neural networks, e.g. the ART networks [46]-[48], between "long term memory" (stored as the strength of the connection weights) and "short term memory" (patterns of activation).

Subtle as it is, this distinction between computer memory and neural network memory underlies an important functional distinction. In the digital computer, stored data must be addressed. It is true that many computers contain a small amount of data storage (content-addressable memory) that is accessed by partial matching of the stored contents rather than by a "memory address"; in both cases, though, access to the stored data is made by means of other data representations (themselves either stored, computed, or presented as part of the data in a coded instruction). Data access is therefore designed around known relationships whereby the information about where the particular operational data required is located is either explicit or calculable from explicitly located data. The means for this is termed the machine's "addressing modes." In a neurocomputer, the memory function is "dynamic" in the sense that signal processing *pathways* are altered in a "self-organizing" fashion (i.e., not under the user's direct control and often not directly accessible by the user). If we adopt a rather generous usage of the term, we could say that all "storage" in a neurocomputer is content-addressed, but it would be more proper to say that the memory function is partly "figurative" (perceptual in the case of recognition functions, imitative in the case of reconstruction functions, and imaginative in the case of evocative functions), and partly operative (consisting of dynamically- and self-organized schemes of representation or actions) [56].

C. Logical-Arithmetic Function. From their very beginning digital computers were intended to be machines for automating logical and arithmetic computations. They are therefore designed to incorporate already-developed constructs of logic and arithmetic. Differences in computer architectures with regard to the logical-arithmetic function essentially come down to decisions made as to which functions will be provided directly in machine hardware and which functions will be relegated as tasks for the computer software to perform. [3] presents the earliest example of this sort of reasoning in the design of the digital computer. In a properly designed general purpose digital computer, the combination of hardware facilities and software instructions permits, if the user so desires, the expression of set-theoretic mathematical constructs. It is this ability that provides the digital computer with its ability to express any mathematical operation that can be carried out constructively in a finite number of steps. This is also the characteristic of the digital computer responsible for the accumulation of errors in its calculations, a characteristic that von Neumann described in terms of the "logical depth" and the "arithmetical depth" of operations performed serially [2].

Von Neumann was convinced that the brain, like the computer, must have both an arithmetical as well as a logical computational capability. However, he was quick to point out that the nature of logico-arithmetic functions of the brain were clearly of an entirely different character from those of formal mathematics and formal logic. As he put it, "the language of the brain is not the language of mathematics" [2]. He formed no conclusions as to how and in what manner the brain expressed logico-mathematical operations, but he did express the opinion that however these operations are carried out in the brain, they were such as to include significantly less logical and arithmetical depth than is encountered in formal logic and mathematics. He proposed, as an hypothesis, that the logico-mathematical functions of the brain were *statistical*, and that the brain's "system of notation" was expressed in periodic or nearly periodic trains of impulses (action potentials). He was the first to suggest that information in the brain was rate-encoded, and that it was "perfectly plausible" that statistical relationships among multiple, simultaneous impulse trains also transmit information. Von Neumann's "statistical hypothesis" underlies many of the most important developments and techniques used in artificial neural networks to this day.

It is a mild presupposition to guess that the logical-arithmetic function in a neurocomputer is most likely distributed among multiple, probably small, neurocalculator networks. Indeed, this is an hypothesis put forward and argued for by Minsky and Papert [57], and one which they view as necessary for the practical solution of a number of issues involved in the scaling problem for learning rates [58], the information representation issue in large-scale networks [57], and the reliability issue often encountered in scaling up small-network solutions to larger but generically identical structures [57]. If this hypothesis is sound, then it follows that techniques and methods learned from research on those neural network structures that this paper calls neurocalculators are of significant importance in the design of general purpose neurocomputers. It also suggests that the solution to the above-mentioned issues with large-scale neural networks might be addressable at the *architecture level* of the design rather than at the scale of the neurocalculator networks themselves.

Nonetheless, the logical-arithmetic function of a neurocomputer cannot be said to be understood merely on the basis of qualitative arguments such as those just given. Instead, it is necessary to take a hard look at what sort of minimally necessary functional capabilities must be present in a neurocomputer. This is not an easy task to carry out in a formal rather than in an *ad hoc* fashion, but one approach, which seems to this author to hold great promise, is to examine the underlying structure of mathematics from the viewpoint championed by the Bourbaki mathematicians. This viewpoint is made all the more interesting to the topic of neurocomputer architecture because there is an amazing parallel between the formal theory of the Bourbaki and the findings of experimental developmental psychology [20]. The Bourbaki mathematicians proposed three "mother structures" from which all other mathematical structures can be generated, and which are not reducible one to another. These are: algebraic structure, order structure, and topological structure.

The prototype representative of Bourbaki algebraic structure is the mathematical group. By way of review, a group is a structure consisting of a set of identifiable elements (which implies that we have a means of element classification) along with an operator, $\circ$, that has the following properties: closure, associative property, the existence of an identity element, e, such that for every element a in the set $e \circ a = a \circ e = a$, and the existence of inverses, a^{-1} , for every element a in the set such that $a \circ a^{-1} = a^{-1} \circ a = e$. A group is constructed from simpler structures, beginning merely with a set of elements and an operation that exhibits closure (a groupoid), and successively adding the other properties to obtain a semigroup (a groupoid with the associative property), then a monoid (a semigroup with identity), and finally a group (a monoid with unique inverses for all elements).

The Bourbaki order structure is a structure of relationships that applies essentially to classes, and its distinctive characteristic is the ability to identify variations that differentiate classes and to set up systems of reciprocal relationships. Its prototype is the lattice structure, and the form of reversibility of an order structure is reciprocity. A lattice is a partially ordered set (a "poset") where every pair of elements in the set has both a least upper bound and a greatest lower bound. The ordering relation in a poset is reflexive, antisymmetric, and transitive. Reciprocity in an order structure means that if $\leq$ denotes an ordering relation such that $a \leq b$ then the reciprocal relation is another ordering relation $\geq$ such that $a \leq b$ implies $b \geq a$. (The two relations symbolized here as $\leq$ and $\geq$ do not necessarily denote the familiar "less than or equal to" and "greater than or equal to" relations).

The Bourbaki topological structure is based on the ideas of neighborhoods, borders, and the approach to limits [86]. A topological structure is a set of elements, E, and a rule, τ, for assigning a system of neighborhoods for every element in the set. τ is called the topology and (E, τ) is called the topological space. The neighborhood of an element x is a subset of E that contains an open set U such that $x \in U$.

Now, what makes the Bourbaki "mother structures" very interesting is that from developmental psychology there are experimental findings pointing to the spontaneous development of fundamental psychological structures that strongly resemble these mathematical structures. **A structure in this psychological sense is a system of self-regulating transformations, these transformations being distinct from the properties of the elements of the system, such that: 1) the result of any transformation remains within the system (closure); 2) in order to apply any of the transformations it is not necessary to go outside the system to find some external element; and 3) the system may have or produce sub-systems through differentiations of the transformations, and can have transformations from one sub-system to another** [20], [60]. Furthermore, it is found that there exists a small set of apparently innate constitutive psychological functions, called coordinators [61], and a small set of compensation behaviors [62] that, taken together, are sufficient to realize these psychological forms of the Bourbaki "mother structures." In addition, it is known that human infants' first perceptual abilities are elementary topological features, and that the subsequent development of spatial, causal, and geometrical relationships is built up from these innate topological feature perceptions [51], [63]-[64]. Finally, it is known that human logical thinking is an outgrowth of a much more primitive "logic of actions" [53], and that the development of logical and arithmetical thinking [65] also follows from these psychological "Bourbaki foundations." Taken all together, these findings of empirical psychology point us in the direction of regarding the neurocomputer logical-arithmetic function in terms of these three logico-mathematical Bourbaki structures, and especially in terms of the primitive topological, coordinator, and compensation behavior functions that produce them.

These psychological findings have an important implication for general purpose cognitive neurocomputation. Rather than coming equipped with a "built-in" arithmetic-logic unit, human-like cognitive systems must come with a "built in" *process* for *structuring* its own logical and mathematical maxims of thinking ("thinking" in the sense that will be defined below). The "innate" capabilities required are the minimal Bourbaki mother structural *forms* and the capacity through *constitutive* coordinator functions and compensation behaviors to build up from these Bourbaki structures specific *constituted* logico-mathematical functions. Thus, what must be provided for are not specific logico-mathematical functions but rather forms of logico-mathematical functionals. In addition to this, the system must come with some form of "built-in" decision-making capability by means of which specific logico-mathematical schemes it constructs are established. This "reasoning and decision-making" capability must at the same time serve to limit the number of potential constituted functions that must be "investigated" or else the system's learning ability will be increasingly compromised as its base of acquired "knowledge of experience" grows (an issue similar to what Bellman termed the "curse of dimensionality"). Now this type of "optimization" problem is already well known in the field of optimal control [85]. Its solution follows **Bellman's principle:**

> An optimal policy has the property that no matter what the previous decisions have been, the remaining decisions must constitute an optimal policy with regard to the state resulting from those previous decisions.

Mathematically, this is usually expressed by the well-known Hamilton-Jacobi-Bellman equation [85]. From the statement of the Bellman principle it follows that this type of optimization constitutes a strategy that works "backwards in time" from a goal to a condition. In terms of Woods' model this corresponds to the EXPECT functionality, and in terms of Piagetian psychology it corresponds to Piaget's Principle of the Central Process of Equilibration (which is discussed later in the section on "reason"). In a nutshell, the logical-arithmetic function of a "cognitive" system is self-organizing and what the system must "come equipped with" are "mother structure" forms and a viable learning process capable of both: 1) the establishment of "goals" and 2) an adaptation strategy that works backwards from goal to decision. Later in this paper, this capability is called "speculative reason," and this brings us to the topic of control function.

D. Control Function. From the earliest conception of the stored-program digital computer, it was obvious that the computer would have to incorporate a control function, the task of which was to cause the proper sequential execution of the operations available within the machine's structure. Part of this task is to translate the coded machine instructions into specific and properly-timed operations. Part of this task is to control access to the computer memory function. Part of it is to ensure the proper employment of the logical-arithmetic function. And part of it is to ensure proper operation of the input and output functions. Machine instructions are said to be "meaningful" only in the sense that a particular instruction code designates a specific sequence of operations.

The meaning and role of a control function is significantly less obvious when the machine under consideration is one which operates in the absence of explicit step-by-step programming instructions, such as is the case for "intelligent agents" or "cognitive systems" or for a neurocomputer. Such a machine, it is true, may have external "control inputs" that signal the machine to undertake particular types of operations. It might even be equipped with a facility by which externally-generated instructions might be given it, a capability which might be called a "language" in some sense. However, it is equally valid to regard such inputs as these as constituting part of the external stimuli that affect machine operation, hence to regard these constituents as part of the input function capacity rather than as a control function in the computer architecture sense of that term. How the machine responds to such stimuli could then be regarded as a manifestation of "behavior" in the psychological sense of that term, and the "behavioral characteristics" it then displays would be consequences of its control function.

In a machine with activities and operations that are said to be "self-organizing" the mechanism that produces the rules of transformation that effect self-organizing behavior is an obvious candidate for the title of the control function of the machine. In the field of artificial intelligence, systems such as theorem provers or "expert systems" employ pre-written software modules as the means of mechanizing the self-organizing rules of the system. Such modules are known by various names, including "meta-knowledge", "meta-logic", and "meta-rules" [66]. The design of such mechanisms is closely tied to various specialized logics including modal logics and temporal logics [67]. Connectionist-based neural network schemes incorporating a "knowledge base" function as a faculty of rules for producing decisions and actions have also been built [68]. One finds in systems such as these particular functions that go by such names as "inference engines" and "rule-based reasoning and learning modules." Since such functions do serve to control the operations of other parts of the system, they qualify as control functions within the context stated above. On the other hand, it would likewise seem that the adaptation rules by which the weights in a neural network are altered, e.g. the well-known backpropagation algorithm, can equally well be regarded as mechanisms of self-organization in a neural network, therefore as control functions. But if we adopt this view, then it is difficult to see how other types of mechanisms, such as the feedback circuitry of an automatic gain control system for a simple amplifier, would not equally qualify as control functions. Yet such a function is more commonly regarded as part of the signal processing in the system in which it is embedded, and it becomes increasingly objectionable to regard such simple functionalities as belonging to the same class of functions as those we intend to designate by the term "control function."

Lacking a criterion somewhat more specific than is provided by the phrase "self-organizing transformation," the number and diversity of the aforementioned mechanisms tells us that this view of what constitutes the control function of a neurocomputer lacks the crispness and clarity which that term holds in digital computer architecture. To put it in other words, what *is it in* mechanisms such as those mentioned above that *constitutes* the logical essence of "control"? Here it is helpful to draw a distinction between functions that are *constitutive* and functions that are *regulatory*. Constitutive functions are functions that *determine structures and operations*. Regulative functions are functions that *determine the employment of constitutive functions*. Regulative functions have two critical roles within a system: 1) the regulations they produce act to preserve existing structures; and 2) the regulations they produce control the modification or enrichment of structures during adaptation [69].

Now, in an open system, i.e. a system that undergoes modification and change in response to outside stimuli that act as aliments of change, the system's regulations need not be fixed nor does their number need be conserved. In other words, regulations can themselves be produced as a consequence of rules of self-transformation. What is more, regulations can be produced hierarchically, i.e. there can arise regulations of regulations. Indeed, such a characteristic seems to be necessary in an open structure if it is to be possible within the structure to produce sub-systems and then to link these sub-systems by rules of transformation. But at the same time, the existence of regulations implies the presence of a regulator, and at some level within the system there must be some form of fixed "master regulation" to control this process (or else we find ourselves in an infinite regression). Such a "master regulatory function" can be called a *functional invariant*, and in both organic systems and in the evolution of intelligence in human children two such functional invariants can be identified: the functional invariant of *organization* and the functional invariant of *adaptation* [70]. Organization refers to the function of organizing the totality of the system in an overall structure, while adaptation refers to the process by which the system assimilates environmental elements into the overall structure and, at the same time, accommodates the system in such a way as to make possible that assimilation while preserving the integrity of the structure as a whole. As Piaget put it, **adaptation is an equilibrium between assimilation and accommodation** [70].

In this context, organization and adaptation are two sides of the same coin, the former denoting the internal aspects of the system for which the latter is the behavioral or "external" aspect. Organization and adaptation, as well as assimilation and accommodation, are easy enough to spot in living organisms. One of the most important and fundamental of Piaget's many findings is that these functional invariants also operate in the development of intelligence and thought in children. The central entity by which organization and adaptation operate is the **"scheme," which is defined as "any activity that can be repeated and generalized."** Piaget discovered that the development of, and probably the formation of, knowledge takes place by means of a *central process of equilibration* which leads from certain states of equilibrium to other, more robust states of equilibrium by passing through multiple "non-balances" (states of incoherence) and re-equilibrations. Equilibrium is achieved by adaptation, and **a state of equilibrium can be defined as a closed cycle of activity in which no further innovations take place**. A non-balance can correspondingly be described as one in which activity does produce innovations [71].[3] The process of equilibration works by means of the mechanisms of assimilation and accommodation. These are governed by two postulates:

1. Any scheme of assimilation tends to incorporate all outside elements compatible with its nature into itself;
2. The entire scheme of assimilation must alter as it accommodates to the elements it assimilates, i.e. it modifies itself in relation to the particularities of events but it does not lose its continuity, and can therefore maintain closure and function as a cycle of interdependent processes, nor does it lose its earlier powers of assimilation.

Assimilation is generalizing, i.e. it incorporates new elements into the common structure of the scheme. Accommodation is specifying, i.e. it leads to the differentiation of structured sub-systems within the general scheme it accommodates. However, these sub-structures remain part of the general structure; they follow the rules of transformation that apply to the general structure, but they can also have specific rules of transformation that apply strictly to the sub-structure (specifications), and they can form rules of transformation that link sub-structures to one another within the general scheme of activity.

There are, of course, many additional details describing the specifics of this central process of equilibration. These are described in [62] with additional important interpretations presented in [52]-[53]. Adaptation is equilibrium between assimilation and accommodation, and a cycle of adaptation either succeeds in assimilation of new elements (while accommodating the structure of the scheme into which these elements are being assimilated) or else the cycle of adaptation *ruptures*, which marks a failure to assimilate [70]. It is the proposition of this paper that **the mechanisms that regulate the central process of equilibration in the neurocomputer comprise the control function.** It is noteworthy in this regard that the central process of equilibration is not immediately concerned with the generation of representations of objects and entities, although these representations are in fact produced as a by-product of this process. The central process of equilibration is the master regulator of the system, and it cares no more for what particular objects acquire representation within the neural system than the control unit of a digital computer cares what particular number is presented to an accumulator by the arithmetic/logic unit. The only goal of the process of equilibration is equilibrium for its own sake. ***Equilibration perfects structure***.

III. Time-Locked, Multi-Regional Neural Networks

The most important question that follows from the discussion in Section II is: What form of neural network system provides a viable foundation for the implementation of the general considerations discussed above? In this paper it is proposed that a general neural network organization first proposed by noted neurologist A. Damasio [45], [49] fulfils these requirements.

A. Pulse-Coded Neural Networks. It is well known that biological nervous systems carry out inter-cell signaling by means of both pulsed signals (action potentials, APs) and non-pulsed amplitude signals. The former type of signaling is mediated by chemical synapses while the latter is mediated by electrical gap-junction synapses. It is a reasonable hypothesis that information is somehow or other carried by these signals, i.e. that these signals are to be regarded as information-bearing waveforms. It is also well known that the same information can be given different *data representations* (if we allow that "signals" can likewise be regarded as a data form), and that, properly done, each of the different possible data representations are equivalent in terms of their information-content [22]-[23]. The key to doing so is know what constitutes "properly," and understanding this is tantamount to answering the question: What is a sufficient neural code for the network? In computational neuroscience the question of the neural code is one of the key unsolved problems. In engineered neural networks, however, we have the design choice of the manner in which information is to transmitted from one neuron to the next. Data communication theory tells us that we have four "pure" choices of modulation schemes for information-bearing waveforms:

[3] The idea of "innovations" is mathematically exemplified in Kalman filter theory.

amplitude modulation (AM); pulse rate modulation (PRM), also known as pulse frequency modulation (PFM); pulse width modulation (PWM); and pulse position modulation or "delay modulation" (PPM). We may also use some combination of these modulation methods, and we may use different methods within different sub-networks in the system. The choice or choices we make impacts the size, cost, and reliability of the design [24].

Software-based GCNNs and Chua's hardware-based cellular neural network (CNN) [25] opt for the AM representation scheme. Information is carried in the amplitude of an analog output signal from each neuron. In the case of the GCNN, the original idea was that the amplitude of the output signal was to be regarded as analogous to the firing rate of biological neurons, the presupposition being that biological neural networks employed PRM as the neural code. This is an idea that goes back to von Neumann's original work on the brain vs. computer analogy [2]. It is an approach that leads to more efficient software models for neural networks, but it is a choice that has a number of serious cost, area, power, and reliability issues for very large scale integration (VLSI) of circuits in hardware [10], [24]. These issues can be ameliorated by the use of pulse-mode artificial neurons, which have been shown to be capable of performing the same mathematical operations as those which describe GCNN dynamics [26].

There are a variety of ways in which pulse-coded neurons can be realized in VLSI. Murray et al. [10], [24], [26] use a charge-transfer method in combination with a voltage-controlled oscillator (VCO). Liu and Frenzel use a simpler and more direct method employing a leaky integrator (LI) charge storage node and modulated ring oscillator [27]. Wells et al. employ an integrate-and-fire technique in a circuit capable of both PRM and PWM modes of operation [28]-[32]. Johnson et al. proposed a more complex pulse-coupled neuron employing multiple leaky integrator circuits combined with a pulse generator circuit [33]-[35], and VLSI implementations have been reported [91]-[92]. It has been demonstrated that pulse-coded neural networks (PCNNs) are capable of carrying out a number of important image processing functions including segmentation, rotation, scaling, and translation [36]. Other pulse-coupled neurons and PCNN schemes have also been proposed [37]-[38].

Although PCNNs have demonstrated very versatile capabilities, particularly for image processing applications, they are by no means free from important implementation issues. Three issues in particular stand out as critical problems for VLSI implementation of PCNN neurocomputers. The first is the weight storage problem [10]. Any neuron that has non-fixed, adaptive weights must have some means for storing the weight-setting information in each adaptive synapse. This is a serious issue because most neurons in a network have multiple synapses, and if a weight must be stored it increases the size of the synapse. The chip area taken up by a neuron's synapses is already the main determining factor of the area requirement of the neuron. One way to store the synaptic weight setting is to store it as an analog voltage on a capacitor. This is, in effect, a form of dynamic memory storage and the method typically will require some sort of dynamic refresh circuitry in order to preserve the weight value [24]. Another method that has been proposed is to use amorphous silicon [24] or floating gate [90] CMOS techniques to implement static weight storage at each synapse. A third method is to use storage registers to set the synaptic current evoked by AP inputs [27]. Regardless of the method used, implementation of settable synaptic weights will generally increase the size of the synaptic circuitry, thereby reducing the number of neurons that can be integrated in a single chip. There is, however, another possibility vested in the *architecture* of the PCNN. This alternative is to construct networks that are capable of performing their computational task by means of a network topology in which *most neurons can employ fixed weights* and only a relative few neurons actually require adaptable weight settings. The feasibility of this method is discussed below.

A second serious issue is that of inter-chip communications. While the issue of on-chip connections is serious enough in its own right for neural networks, this issue is compounded when a neurocomputer requires a multi-chip design. At the present state of the art, large-scale neurocomputing will almost certainly require multi-chip implementations. Here the key issue is the number of interconnections required to link one PCNN chip to another. The fundamental limitation is packaging technology. Chip-to-chip interconnects requiring a large number of interneuronal signals likewise requires a large number of input pins be available on each package, and a similarly large number of output pins. Even with exotic packaging technologies, this issue stands as probably the main impediment for the implementation of large neurocomputers [24]. One possible approach to this problem is to implement some sort of multiplexing technique so that a single interconnecting wire can carry signals from multiple neurons [24]. This, of course, requires that multiplexing and demultiplexing circuitry be included in every chip, along with the necessary control and timing signals for implementing reliable communications between chips. Another alternative is again the architectural alternative of coming up with a network topology that minimizes the amount of inter-chip communication required. This, too, is discussed below.

The third issue is the learning strategy employed by the network. Although fixed-weight PCNN solutions have been shown to be viable for specialized computing tasks, e.g. image processing, it is clear to everyone that a general-purpose hardware neurocomputer *must* have some form of learning capability if it is to be able to deal with unpredictable and changeable operating environments in its applications. To date few learning algorithms for hardware PCNNs have been published, although it is widely assumed that some form of Hebbian learning strategy is a likely candidate at the neuron level [24]. A complex scheme based on the Informax Principle has also been proposed [93] Pulse-mode neurons are not particularly well-suited to standard learning algorithms such as gradient descent, and a PCNN is not well suited for standard network learning strategies such as backpropagation or the MADALINE rules [39]. At the same time, known good Hebbian algorithms such as the Bienenstock-Cooper-Monroe (BCM) family of algorithms [8] are likely to prove too prohibitively expensive to implement in each neuron in the network. Obviously, this issue is also tied to the weight-storage issue discussed earlier. Less obvious, but a proposition argued in this paper, is that the learning strategy of the neurocomputer and the architecture of the network are coupled issues, and need to be *jointly* considered in the design of a neurocomputer.

B. Time-locked neural coding. It is widely agreed among neuroscientists that there probably is not one unique manner of information representation in biological neural networks. Sensory neurons for the most part appear to encode information in their firing rate, a PRM scheme. It is much less clear how more complex objects and entities are represented within biological neural networks. Theorists have proposed schemes ranging from that of the so-called "grandmother cell" scheme (the method most often employed in artificial neural networks) all the way to the "distributed representation" or "vector" scheme [40]. In the case of distributed representation, there arises the accompanying problem of how fragmented representations of the features of a complex object are to be connected together. This is known as "the binding problem."

There is good reason to think that neither scheme is actually used in "pure" form by biological neural networks to represent entities and events. The choice of data representation scheme for entities and events is one of the most crucial design decisions that must be made for a neurocomputer architecture, and if there is merit in using "biologically inspired" schemes for modeling individual neurons, there is even more merit in using "biological inspiration" in the case of the neural coding representation.

Over the course of the last few years, there has been growing neuroscience evidence pointing to the hypothesis that complex object representation in biological networks is based on time-locked, synchronous firing activities of entire assemblies of neurons. Malsburg and Schneider were among the first to give prominence to the idea of time-locked firing patterns as the basis of segmentation and object representation in neural networks [41]. Not long afterwards Eckhorn et al. proposed the use of time-locked firing patterns as a possible solution for the binding problem [42], and the Eckhorn model was quickly adopted for use in PCNN image processing by Johnson and others. Some starting-point foundations for the mathematical description of time-locked system dynamics were provided by Senn et al. [43], although a thorough understanding of time-lock dynamics in the form of a mathematical theory is yet to be achieved. Even so, one thing is already quite clear about time-locked PRM signaling: It provides in principle a straight-forward method for evaluation of information loss characteristics for signals converging at any particular neuron.

Information loss in computing is not a bad thing provided the properties of that information loss are well controlled. For example, it is a trivial proof in information theory to show that all arithmetic and logical binary functions are information-lossy. If I have an adder performing unsigned arithmetic and its output is 5, I do not know if its inputs were (5, 0), (4, 1), (3, 2), etc. Because knowledge of the output alone does not let me unambiguously determine what the inputs were, the adder is information-lossy (as that term is used in information theory). Provided that an input pair such as (2, 2) does not also produce 5 as an output, this information loss is not only harmless but, indeed, it is *necessary* for the proper functioning of the adder. This example is analogous to the well-known neural network processing operation by which an input space is divided into regions according to the decision boundaries established by the network.

More formally, let X be the set of possible input "symbols" that can be applied to a processor and let Y be the set of output symbols produced in response to these inputs. The amount of information contained in the input set is its entropy, $H(X)$, and the mutual information between input and output, which is a measure of how much our uncertainty of one symbol is reduced by knowledge of another symbol, is [22]

$$I(X;Y) = H(X) - H(X|Y) \tag{1}$$

where $H(X|Y)$, $0 \le H(X|Y) \le H(X)$, is called the equivocation and is a measure of how much uncertainty there is remaining in X given knowledge of Y. The processor is information-lossless if and only if $H(X|Y) = 0$. Neural network classifiers work because of precisely-determined equivocation characteristics produced by their weights.

Now let us suppose a neuron has N synaptic inputs and is set up to respond to synchronous, in-phase, time-locked AP input trains of some firing rate $R_i = 1/T$, duty cycle D, and zero relative phase delay. Let us further assume an integrate-and-fire neuron model, and that the weighted sum of n action potentials inputs (n active time-locked input firing trains) applied to the integrator has a value α. Letting the firing threshold of the neuron be θ, it is easy to show that the output firing rate is a frequency-divided version of the input firing rate

$$R_o = \frac{R_i}{d} = \frac{R_i}{\left\lceil -\tau R_i \ln\left[1 - \frac{1-\exp(-T/\tau)}{1-\exp(-DT/\tau)}\frac{\theta}{\alpha}\right]\right\rceil} \tag{2}$$

where τ is the time constant of the integrator and

$$\alpha > \frac{1-\exp(-T/\tau)}{1-\exp(-DT/\tau)}\theta \tag{3}$$

Note that the input firing rate is an integer multiple d of the output firing rate. For some given set X of time-locked input firing patterns characterized by various rates R_i and duty cycles D, in principle it is possible to find a set of synaptic weights W such that different α terms corresponding to different input signal patterns results in a one-to-one mapping between output firing rates and particular subsets of X. This of course constitutes a set of decision boundaries in the input space and corresponds to the weight-dependent decision boundaries set up in multi-threshold generic connectionist neurons [44].

C. Time-locked multi-region networks. The time-locked representation proposals just cited were applied to neural networks on a relatively small scale. However, there is considerable mounting evidence that the human central nervous system employs this scheme on a vast scale. On the basis of numerous neurological studies, Damasio first proposed that time-locked, multi-regional signaling is the mechanism used in the brain to represent entities and events, and to solve the binding problem [45]. The central idea in Damasio's hypothesis is that binding and synchronization of networks is carried out by small assemblies of cells that he called "convergence zones" (CZs). The firing patterns emanating from convergence zones are called "dispositional patterns" because their main function is *to implement control over multiple, large regions of cell groups and assemblies*. Except for certain features found in ARTMAP networks (i.e., the "vigilance" function) [46]-[48], this is one of the first times that an explicit use has been made of neural sub-networks to implement signal processing *control* functions within a neural network architecture.

Damasio's basic scheme is illustrated in simplified form by figure 2 [49]. This diagram is intended to illustrate only the novel aspects of Damasio's hypothesis and omits many neuronal assemblies at the higher orders of organization, e.g. connections to the motor cortices. Damasio's hypothesis and the neural architecture he proposes have sweeping implications for the design of intelligent systems because it provides a theory of organization for neural substrates not merely for specialized information processing tasks (e.g. image processing), but also for the entire spectrum of cognitive, learning, and memory functions. Convergence zone neural assemblies are not feature-detecting integrators of lower order information; rather, they function as "pivots" inscribing amodal records of synchronous combinatorial arrangements of representations from upstream networks. Each CZ provides one-to-many feedback signals to immediately upstream networks in the form of one of two types of "binding codes" that give rise to integrated representations of entities and events represented by the distributed activity patterns of upstream networks. The system is comprised of four main classes of neuroanatomical substrates [45]:

1. Feature networks – primary and early sensorimotor association cortices whose representations are formed into representations of entities by Type I binding codes, and into representations of events by Type II binding codes;

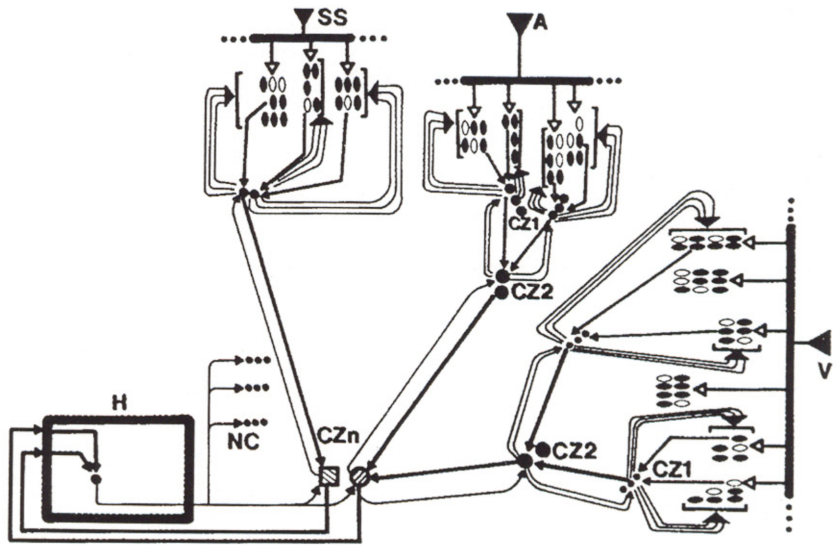


Figure 2: Simplified diagram of time-locked, multi-regional neural network architecture. The figure is taken from Damasio [49]. V, A, and SS denote early and intermediate sensory cortices in visual, auditory, and somatosensory modalities. CZ denotes convergence zone cell assemblies. H denotes the hippocampal system. NC denotes non-cortical neural stations in basal forebrain, brain stem, and neurotransmitter nuclei. Convergence zones are organized as a layered system of several orders (1, 2, ..., n). Dark lines depict feedforward connection paths, light lines depict feedback pathways. The figure is not a complete depiction, e.g. it omits the details of connections made to the output section of the motor control networks. The structure is a "layers of networks" rather than a "layers of neurons" organization.

2. convergence zone assemblies – small networks of neuron assemblies in sensory and motor association cortices of different orders, some limbic structures (entorhinal cortex, hippocampus, amygdala, cingulate cortices), and the neostriatum/cerebellum;

3. feedforward and feedback connections. Feedback connections project back to those upstream sites making immediate connection to the convergence zones, regardless of whether these upstream sites are feature networks or other convergence zones of immediately lower order. The connections in this structure are to be viewed as connections made to lower order subnetworks rather than as connections back to those neurons within these networks that project immediately to the convergence zone assemblies as output neurons;

4. non-specific thalamic nuclei, hypothalamus, basal forebrain, and brain stem nuclei. These assemblies function as providers of indirect feedback pathways *that modulate the activities* of large regions of neural groups; they are also involved in the initial *formation* of convergence zones.

In contrast to the traditional neural signal processing model, in which successively downstream neurons act as integrators and feature recognizers said to "store" or "recognize" or "classify" representations of entities or events, the system proposed by Damasio works by stimulating time-locked co-activation of multiple subnetwork regions within the sensory and motor cortices. Each of these regions provides some "fragment" or "feature" of an entity or event, and it is the totality of this stimulated neural activity that constitutes the representation of the entity or event. There is substantial neurological evidence which supports Damasio's hypothesis and which at the same time contradicts the traditional model [45]. The information processing that takes place in the network is both parallel *and* sequential, the latter because of the many time phases actuated in multiple steps by feedback from the convergence zones which binds the various regional signaling activities into a coherent whole of representation. Damasio's hypothesis holds that: 1) perceptual experience depends on neural activity in multiple regions activated simultaneously, rather than on a single region where experiential integration occurs; and 2) during both free recall and recall generated by perception in a recognition task, the multiple regional activity necessary for experience occurs near the sensory portals and motor output sites of the system rather than at the end of an integrative processing cascade removed from the system inputs and outputs. The process of re-stimulation of selected features in the early networks processing sensory modalities can, without philosophical prejudice, be called a "synthesis of imagination" since the activities so reproduced correspond to an integrated representation of entities and/or events. Co-activation of motor outputs can likewise be called a "motoregulatory expression" [50] since in living organisms the coordination of motor activity with sensory input processing is known to play a fundamental role in the synthesis of cognition [51], in the development of cognizance [52], and in the establishment of "meanings" of perceptual representations [53].

In this architecture, the patterns of neural activity that correspond to distinct properties of entities are expressed in the same neural ensembles in which they occur during perception, but the binding codes that synthesize their combinatorial arrangements into entities or events are recorded in the convergence zones (CZs) [49]. Feedback of binding code activity patterns takes place through three distinct routes:

1. Direct routes, where CZ signals travel directly back to the feeding networks;

2. indirect routes whereby feedback from higher-order CZs reaches the early sensory networks via intervening lower-order CZs;

3. non-specific routes via metabotropic signaling pathways [30] that influence widespread regions and/or the system as a whole via *modulatory* signals that alter the basic signal processing states of subnetworks in the system. Such feedback pathways can alter the excitability of subnetworks and can even enable or disable the active connections among neurons and/or neuronal assemblies, thereby implementing a switched variable structure in the basic network organization. This is illustrated by a biological example shown in figure 3 [54].

In addition to these feedback pathways, CZs project forward to downstream networks, possibly including additional feature-representing networks. Cross-projection to other assemblies at the same level is also possible, either by direct lateral coupling or indirectly via the one-to-many feedback pathways to upstream networks [45].

Convergence zones that bind features into entities are located in the early portions of the signal processing stream, i.e. they are lower-order CZs. Convergence zones that bind entities into progressively more complex events or entities are distributed along the higher-order segments of the pathway, i.e. they are higher-order CZs. Damasio defines two types of convergence zones. Type I zones fire back simultaneously, inscribe temporal coincidences and aim to reproduce them. They are located in sensory association cortices of both low and high order and are assisted in learning by the hippocampal system. Type II zones fire back in *sequences* and produce closely ordered activations in the target cortices. They inscribe temporal sequences (events) and are prominent in motor-related cortices. They are assisted in learning by basal ganglia and the cerebellum. The two types of convergence zones interlock at multiple levels such that learning relative to an entity or event recruits both types.

Damasio's hypothesis places synaptic learning in the convergence zone assemblies. Although his theory is non-specific with regard to structure at the synaptic level, the signaling activity scheme he describes raises the important possibility that *long term weight adjustment may be confined to a relatively small fraction of the total number of neurons in the system*. Convergence zones stimulate activity in other neuron assemblies, and this means that it is quite possible that most of the synaptic actions in many of the system's neurons may be *elastic modulations* rather than quasi-permanent changes in synaptic weight. This is a potentially important consideration for the design of large-scale neurocomputer networks. As noted previously, synapses occupy most of the chip area consumed by a neuron. Adjustable-weight synapses tend to occupy more chip area than fixed-weight synapses, and so if the majority of synapses are fixed-weight this significantly reduces the size of most neurons in the system and provides a significant reduction in the magnitude of the weight storage problem. Elastic modulation mechanisms operating at the whole-cell level of the neuron are economical in the sense that they is not subject to multiplication by the number of synaptic inputs [30].

IV. The Organization of "Mental" Processes

This paper uses the term "mental processes" to mean those activities of the neural network system that produce representations of objects (entities and events), internal machine states (which here are called "affective perceptions" to denote that these representations do not become part of the representation of an object), and

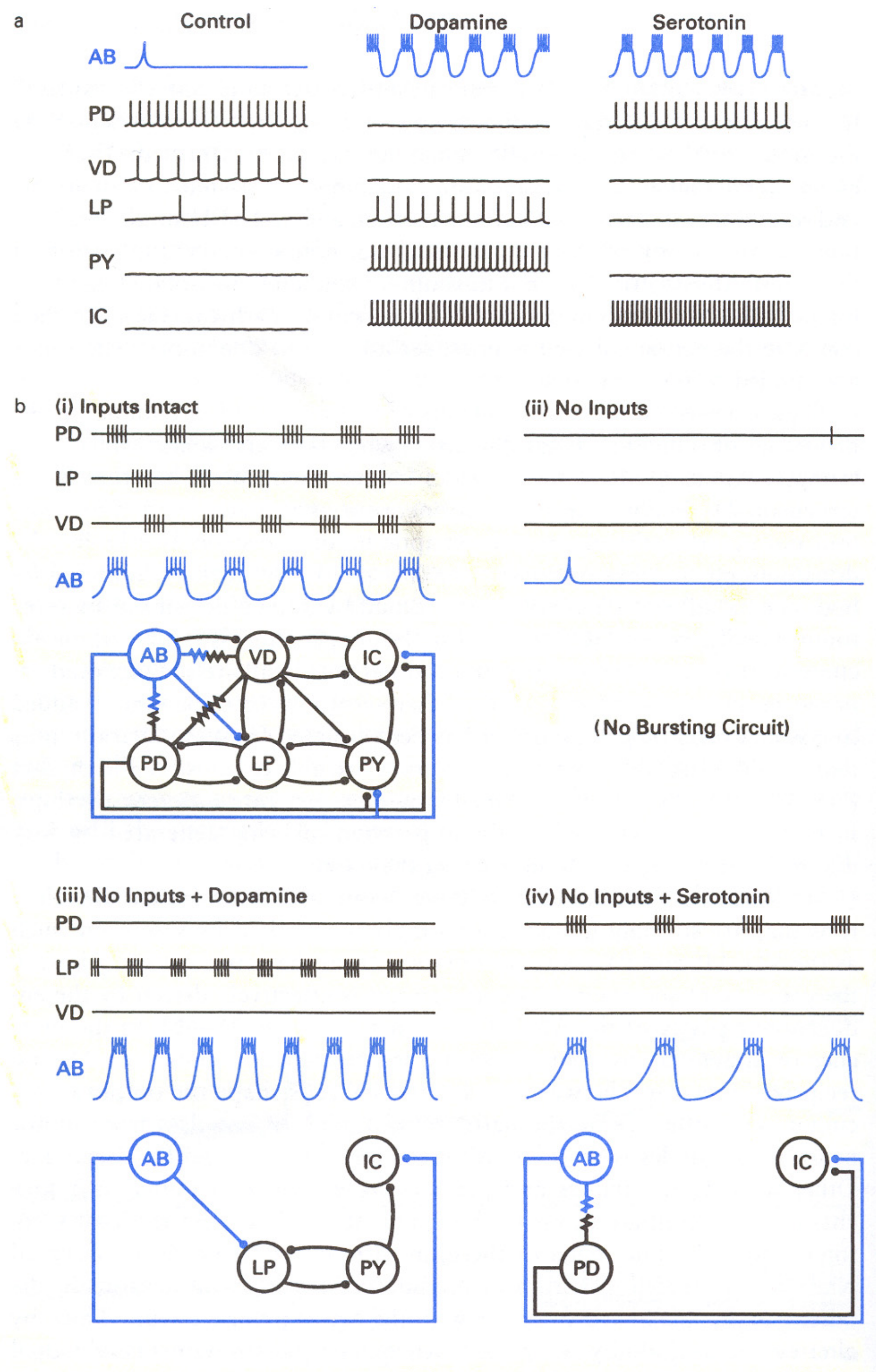


Figure 3: Illustration of neural network modulation via metabotropic signaling. This illustration is a biological example, specifically the pyloric network of the lobster. The figure is taken from [53]. Dopamine and serotonin are metabotropic neurotransmitters [30] and their effect generally is that of control and modulation of signal processing functions rather than data processing. The network illustrated here performs a central pattern generator function. AB, VD, IC, PD, LP, and PY designate particular assemblies or neurons in the pyloric network. Metabotropic inputs to the network re-route its operational signaling pathways.

control/binding codes that direct the operational modes of the system and determine its causal agency. These processes are termed "mental" in order to denote relationships to phenomena that in a human being would be termed the "psychological" co-manifestations of brain function.

The manner of organization of mental processes is, of course, the great unsolved problem of neuropsychology. Psychological objects (concepts, intuitions, emotions, etc.) are not directly observable to neurobiology, and from the viewpoint of ontology they must be regarded, like the term "mass" in physics, as supersensible objects, i.e. as constructs used by science in the construction of theories. This supersensible "nature" of psychological objects has, naturally, spawned a great number of competing theories and an even greater debate as to fundamental definitions of terms such as perception, concept, emotion, etc.

Many of these questions are of a deeply philosophical sort, and since there is today no one philosophical doctrine to which everyone subscribes it is no wonder that there should be such a wide divergence of views on the issues just mentioned. The viewpoint taken in this paper draws its philosophical ground from the work of the great 18th century philosopher Immanuel Kant. Kant made a life-long study of most of the issues that concern us here, culminating in a rich, highly technical doctrine which he initially called "transcendental philosophy" and which he later called the Critical Philosophy [21]. Its starting point is the hypothesis that all our empirical knowledge is of such a nature that it must conform to certain fundamental capacities and processes of the mind. As he put it, objects conform to our cognitions (cognitive powers) rather than the other way around. For Kant there is a distinction between "things" (the existents in nature) and "objects." A Kantian object is *how a thing appears to us*, and all our knowledge about a "thing" is really objective knowledge (knowledge of the thing's appearances). However, this does not mean that we are to doubt the "real existence" of these "things," despite the strong element of idealism in Kant's system. As he put it, "a transcendental idealist is an empirical realist," and how it is not only possible but necessary to be both at the same time is a topic to which he devoted considerable discussion.

Kant's theory gives priority to epistemology over ontology, and he investigated what constituted the necessary conditions required for the possibility of experience and human knowledge, i.e. "knowledge as we know it." Because he took this approach, it was necessary for him to delve into precisely the sort of questions we face here – e.g., "What is perception? What is a concept? What is intuition? What is imagination? What is understanding? What is judgment? What is reason? What is consciousness?" etc. It also led him to identify various processes and functions that must be capacities of the mind if knowledge and experience are to be possible in the way that they are for human beings. These functions he called "knowledge *a priori*," and this form of knowledge is not knowledge of objects but rather a kind of "how-to" knowledge necessary for the possibility of *constructing* objective knowledge and experience in general. The resulting doctrine is magnificent, sublime, highly technical, and self-consistent.[4] It is impossible to even provide a description of his entire theory in this paper; this must be done elsewhere [71]. Therefore, this paper limits itself to a brief overview of those results required for our present topic.

By combining Kant's analysis with Piaget's experimental findings, an organization of mental processes emerges. This organization is depicted in figure 4. The five blocks depicted in this diagram represent various processes operating on information to form various types of knowledge representations, evaluations of the system's state of equilibrium or disequilibrium, equilibration control functions, or motor expressions. This "static" diagram of mental process organization can be compared with Woods' "dynamical" model of figure 1. It is clear that the system depicted in figure 4 is broken out in more detail than is presented in the "reasoning loop" in figure 1. Figure 4 does not depict the logical-arithmetic function because this function is implicit in the formation of Bourbaki structures by the system. The explanation of the processes and information pathways shown in figure 4 is as follows.

[4] There is no shortage of philosophers who will strongly dispute the claim made here that Kant's system is self-consistent. However, the study of Kant in the English-speaking world has always been severely hampered by errors in translation of [21] that distort and often drop Kant's hair-splitting and highly technical distinctions, e.g. existence in the *Dasein* sense vs. existence in the *Existenz* sense, or Kant's distinction between object (*Gegenstand*) and Object (*Objekt*), or his distinction between combination (*Verbindung*), composition (*Zusammensetzung*), and connection (*Verknüpfung*), or the faulty mistranslation of Kant's *Lust und Unlust* into the utterly incorrect "pleasure and pain," to name but a few. Furthermore, Kant's system is not confined to *Critique of Pure Reason* or even to all three of his great Critiques; rather, it is distributed throughout the entire Kantian corpus [21], and one must study the whole to understand its parts. Finally, it helps in the study of Kant's system to have had some training in the discipline of system theory, and this is not common among philosophers. This author finds no internal nor external inconsistencies or contradictions in Kant's system.

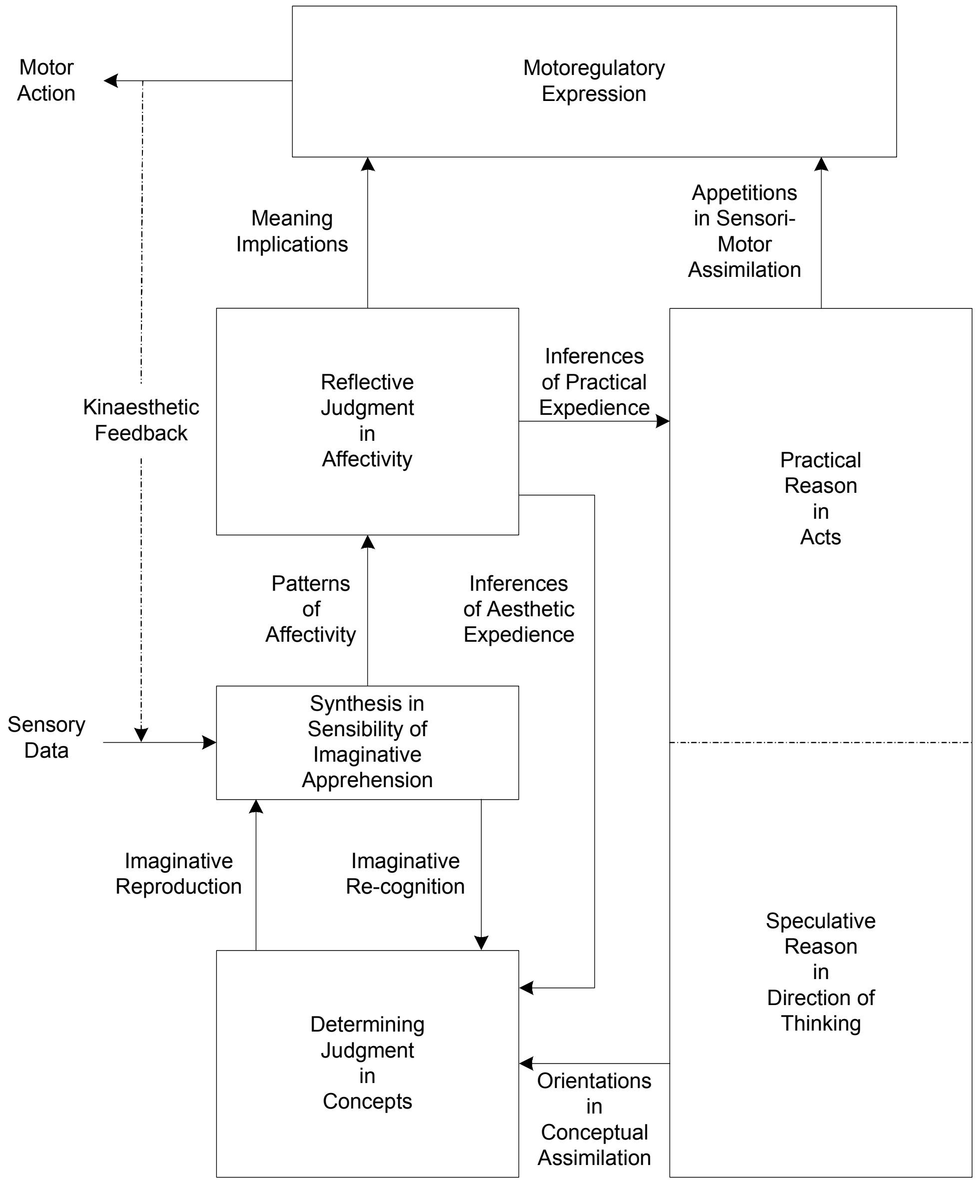


Figure 4: Kantian-Piagetian Architecture of Mental Processes. Blocks show the major processes involved in mental organization. Lines depict pathways for information flow within the architecture. Cognitive functions involve the interplay between sensibility and the process of determining judgment, but also rely on lateral feedback of affectivity codes from the process of reflective judgment. Reflective judgment is non-cognitive and is only concerned with the evaluation of the overall state of equilibrium or disequilibrium present in the system, and with binding codes that link sensibility to motor actions. This linkage is called a meaning implication. The process of reason controls the central process of equilibration. Practical reason is concerned with assimilation in sensorimotor schemes of action; speculative reason is concerned with assimilation insofar as this assimilation is concerned with acts of determining judgment in the assimilation of concepts in conceptual structures. A cognitive representation in sensibility is called an intuition; an affective representation in sensibility is called a feeling. Feelings are regarded as energetic regulations. They do not form structures but they facilitate the formation of structures. Binding codes emanating from the process of determining judgment are regarded as rules for the reproduction of intuitions, and the activity patterns of convergence zones generating these binding codes are called concepts. Binding code activities emanating from the process of practical reason are called appetites and those emanating from the process of speculative reason are called ideas. Binding code activities emanating from the process of reflective judgment are called meanings or inferences of expedience, depending on their target.

A. Sensibility. We will call any representation in neural activity of which the system is “conscious” a perception. “Consciousness” in the sense used here means the system generates a control signal representing that another representation is presented, thereby marking that representation as distinct (salient) from other “background” neural activity. It is this second representation that is the perception. Perceptions are represented by the sensibility process of figure 4, and we will distinguish two types. Objective perceptions will be called *intuitions*. Non-objective perceptions, i.e. perceptions that do not constitute any part of the representation of an object but instead convey knowledge of the operational state of the system, will be called *affective perceptions*.

In terms of the Damasio hypothesis, intuitions are represented by relatively stable, time-locked, multiregional firing patterns from specific feature networks (figure 2) that have been linked by binding code firing activities of convergence zones. Feature networks project to multiple, overlapping convergence zone networks, and so the firing activities of any one particular feature network is to be regarded as a “feature” of an object. Features can be bound to different aggregates of multiregional firing activities and so are, in a sense, “mobile” representations (i.e. can participate in many different object perceptions). Convergence zones generating binding codes for intuitions correspond to the process of determining judgment in figure 4.

Kant’s analysis of sensibility showed that the making of an intuition involves a three-step synthesis that he termed the synthesis of imagination in apprehension. He called the steps in this process *comparation, reflexion,* and *abstraction*. Comparation is a type of comparison operation. It might loosely be called a “likening” operation and more formally be called an operation that establishes a compatibility relation (reflexive and symmetric relations) in and among feature activation patterns. A bit less abstractly, the comparison operation establishes multiregional firing patterns capable of being stabilized as repetitive, sustainable firing activities. If a and b are firing activity patterns from two different feature subnetworks, the comparation operation establishes a binary relation C such that $a\mathrm{C}a$ (reflexive relation, i.e. activity a is a stable activity pattern) and $a\mathrm{C}b \rightarrow b\mathrm{C}a$ (symmetric relation, i.e. a and b can both occur without detriment to each other). In Piagetian terminology, the act of comparation is an association coordinator, i.e. it produces ordered n-tuples of compatible firing patterns.

While comparation can be regarded as a type of “logical” or “mathematical” comparison, reflexion is to be regarded as a kind of “material comparison.” A bit less vaguely, the act of reflexion has to do with the synthesis of a structure in terms of what is to constitute the material makeup of the intuition. For example, there is a great difference between the perception of an object presented immediately to the sensory apparatus of the system and the perception of an *imaginary* object “remembered” without direct stimulation of external sensors. Recalling that Damasio’s convergence zones are capable of stimulating the re-presentation of specific firing activities in the feature networks, but that the original presentation of these firing activities is stimulated from inputs arriving from the system’s sensory apparatus, the *meaning implication* of a firing pattern depends on the originating sensorimotor conditions concurrent with the particular representation of the object. Furthermore, there is an important difference between, for instance, imagining myself as raising my hand vs. actually raising my hand. Whether or not a particular overt *action* will or will not follow upon an intuition depends on whether a connection is established between the activity patterns that represent the intuition and other networks involved in motoregulatory expression. Put more briefly, the act of reflexion establishes connections between firing patterns in feature networks representing an object and other networks that determine how the rest of the system will respond to the representation of the intuition. This means that sensibility produces a third type of binding code, not discussed by Damasio, that can be called a “connection code” that organizes the coordinated activation of different networks in a “network of networks” architecture. Such a code determines “what will be done” with the intuition, and in this sense it is identical to what Piaget and his co-workers called a *meaning implication* [53].

Abstraction can be described as “the segregating of everything by which two representations differ.” However, since activity patterns in sensibility cannot be called “objective” until an intuition is produced, and since the production of an intuition requires the completion of the full three-step process (which includes abstraction), a better explanation than this of the act of abstraction must be set down. Kant said, “Abstraction is the actualization of attention, whereby only a single representation is made clear and all the remaining are obscured.” This statement is easy to understand within the context of the convergence zone theory of neural network architecture. In a recurrent network structure such as we have here, and in a system having sensory inputs that are continuously active and presenting new data to the system, there is obviously a strong potential for activation of a multitude of feature networks such that a veritable “chatter” of firing patterns could be produced. It is clearly necessary for the proper operation of the system that incoherent and possibly conflicting activation patterns be suppressed, and this is the job of abstraction in the synthesis in apprehension. It is just as important to have a mechanism that suppresses firing in sub-networks as it is to have a mechanism for producing activations.

Let us recall that the central control principle in the "psychology" of the neurocomputer is the principle of equilibration. Establishment of stable and *coherent* network activity may in a philosophical sense be called a "pure purpose" of the control function of the system. Note also that this "law" of the system's operation is a law of pure form, i.e. it has to do only with the establishment of conditions of equilibrium and has nothing whatsoever to do with "what" is represented by an intuition. It can in this sense be called an *a priori* law of system operation since it is a condition necessary for the possibility of establishing what we might call "machine cognition" of objects and "machine experience" of objects. Therefore, sensibility must contain a mechanism by which the activity patterns being produced tend toward a state of equilibrium, and this is done by suppressing network activities that run counter to the establishment of such an equilibrium. Damasio called attention "the 'spotlighting' process that generates simultaneous and multiple-site salience and thus permits the emergence of evocations" [45]. Abstraction is the process that makes such a "spotlighting" possible, and its mechanism will consist of neurons in one or more sub-networks that produce inhibition of activity in those feature networks whose outputs are counterproductive to the establishment of stable, coherent activity patterns.

Intuitions are building blocks of higher cognitive structures. However, unless we are to suppose that the system contains either the "innate ideas" of the rationalists or a "copy of reality" mechanism – both of these being hypotheses that experimental psychology finds against – we cannot hold that intuitions "form themselves"; some other type of representation in sensibility must take on the task of providing the information used for *regulating* the formation of intuitions. This is the role of affective perceptions. Affectivity, in and of itself, does not form structures [72]. It is, however, somehow bound up with the phenomenon of cognition. In what way this is so is a matter of open and on-going disagreement among psychologists, with views ranging from, at one extreme, the hypothesis that affectivity follows cognition [73]-[74] to the opposite extreme holding that affectivity precedes cognition [75]. Other researchers, e.g. Piaget [72], Greenspan [76]-[77] and Damasio [78], take an intermediate position in which affectivity and cognition are cooperative yet distinct processes, the combination of which gives rise to consciousness and cognitive structuring. Damasio has proposed that certain neural structures produce affective representations he calls "somatic markers" which play a role in cognitive appraisal and judgment [79].

Piaget proposes that affectivity has a two-sided role. On the one side, affectivity is an energetic regulator of behavior. In this role its relationship to activity and cognition might be likened to the relationship between the gasoline and the motor of a car. Affectivity does not form structures but it plays a role in the regulation of motor and cognitive activities that *do* form structures. These regulations fall into two wide classes: Activating regulations trigger and maintain activities; terminating regulations bring them to a finish.

On the other side, affectivity regulates the evaluation of activities and the construction of a system of values. Piaget describes the term "value" as a "general dimension of affectivity and not a particular or privileged feeling." Somewhat more precisely, a representation is called a value when that representation expresses an indication of a rupture in equilibrium or an incompletion in the attainment of a state of equilibrium [70]. In a general sense, "that which is valued" by the system is a return to a state of equilibrium, and the representation that this has been achieved can be called a "satisfaction" [72].

These two roles of affectivity, energetic regulation and evaluation of activities, therefore have a direct bearing on acts of adaptation by the system. Piaget calls the conjunction of the two roles of affectivity a mechanism of "interest." Interest is the affective aspect of assimilation, understanding is the cognitive aspect of assimilation. Hence, the representation of affective perceptions can be regarded as a measurement (evaluation) of the state of the system viz. its functional expedience for the primary control function of the system (equilibration). For the system model presented in this paper, the condition of equilibrium is the achievement of stable, time-locked activity firing patterns, which implies that the expression of affective perceptions should be based on special neural coherence and incoherence detectors, distributed throughout the system at all levels (but all functionally part of the sensibility process), whose outputs project to networks that determine activation or deactivation conditions of other networks. This is depicted in figure 5.

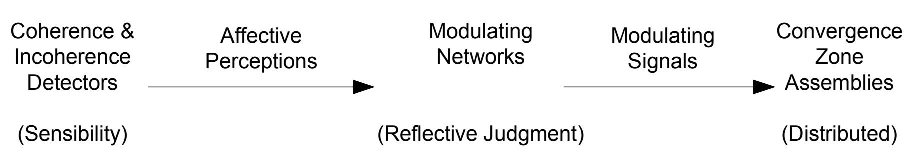


Figure 5: Affective signaling pathway model. Affective perceptions arise in sensibility and project to modulating networks belonging to the reflective judgment function. These networks project modulatory signals to various convergence zone assemblies distributed throughout the system and affect their signal processing modes.

However, since affectivity does not form structures, the hypothesis we make here is that the *direct* conditioning of other networks in response to affective perceptions is primarily by means of elastic modulatory signals. Plastic adaptation of synaptic weights in convergence zone assemblies in some cases might follow indirectly from this conditioning as a consequence of accommodation in signaling structures, but this would be a second-order effect of affective signal processing.

The principle evidence for this hypothesis is provided by neurological studies of the human central nervous system that implicate specific brain structures for affective phenomena. Damasio [78]-[79] holds that limbic system circuitry, particularly the amygdala and anterior cingulate, the hypothalamus, and neurotransmitter nuclei in the brain stem and in the basal forebrain are principal subsystems for activating affective reactions. These systems transmit neuromodulators targeting sub-systems in the cerebral cortex, thalamus, and basal ganglia. It is true that some of these neuromodulators are implicated in long term potentiation and long term depression of synapses, but the more fast-acting effects mediated by these neuromodulators are generally metabotropic signals [30] which can alter active signal pathways (cf. figure 3) and which, from the viewpoint of signal processing, must be regarded as temporary modulations rather than long-term accommodations.

B. Reflective Judgment. Kant defined judgment in general as the subsuming of particular rules under a general rule. There are, however, two ways in which this can proceed. If the general rule is given and particular rules falling under it are to be found, the judgment is called *determining* since the general rule determines the particulars. This is the judicial mechanism depicted in figure 4 as the process of determining judgment. On the other hand, when only particular rules are given and a general rule for them must be found, the judgment is called a *reflective* judgment, and this is the mechanism of judgment depicted in figure 4 as the process of reflective judgment.

Here it is important that we clearly establish what in general is meant by the term “rule.” Kant defined a rule in general as any assertion that takes place under a general condition. Thus, “stop at a red light” is an example of a rule, but Kant’s definition is far broader than this example would suggest. The firing of a neuron in response to specific input conditions can also be viewed as an assertion, i.e. neuron firing activity “asserts” something, just as we say that a logic signal in a computer “asserts” something. The Kantian definition of a “rule” goes well beyond the usual usage of that term in artificial intelligence work, and it gives us a way to regard neural networks as systems that construct and instantiate rules. “Subsumption” under a rule means to represent that the condition of the rule has taken place. In Piagetian terms, a general rule is a structure, and particular rules subsumed under it are sub-systems of that more general structure. In a neural network, a rule is an *a priori* rule if it is initially “wired in” to the structure of the network. If it is “hard-wired” (not subject to accommodation), then the rule might be called a “meta-rule” (fundamental rule about rules) because it forms part of the permanent organization of the system. If it is “soft-wired” (subject to change through accommodation), then it is a rule scheme. If it is not part of the initial “wiring” of the system and is formed through the system’s activities then it is an *a posteriori* rule and its structure is representative of the system’s “experience.” All *a posteriori* rules are “soft-wired” and “empirical.”

Reflective judgment is tasked with finding a general rule for given particulars presented through sensibility. However, in carrying out this task there are a number of special restrictions on this capacity. First of all, empirical general rules are never presented by the data of the senses. Furthermore, Kant (on philosophical grounds) and Piaget (on experimental grounds) both agree that human beings possess no *a priori* “innate ideas” of objects as rationalist philosophers had supposed. Instead, both Kant and Piaget held that apriority with regard to mental capacity involves only certain innate functional abilities that make possible the construction of concepts and ideas of objects. Because general rules are not presented by the data of the senses nor are they present *a priori* as innate *a priori* objects, the process of reflective judgment is a *non-cognitive* process, albeit a process that does regulate the formation of cognitive structures. The task here is to explain how this non-cognitive process leads to cognition through the process of *determining* judgment. There are two main considerations in this, one experimental and arising from research work in developmental psychology, the other theoretical and arising from philosophical grounds.

Kant drew a technical distinction between the terms “nature” and “the world,” and this distinction is for all practical purposes the same as the distinction Woods draws between “world model” and “world.” For Kant, “nature” is the “world model” that the mind constructs in experience. Empirical laws are general rules by which we *make a system* of nature, binding together particular elements of experience in a unity of interlocking relationships. The process of reflective judgment is tasked with finding such general rules, and therefore with the task of making a system of nature in experience. In doing so, Kant deduced that reflective judgment follows an *a priori* principle that he called the principle of formal expedience (*Zweckmäßigkeit*) of nature. The *form* of a representation is expedient if and only if, as a condition for the cognition of an object, its agreement with the experience of that object is possible only according to a purpose. From a judicial standpoint, we can say that the purpose of a reflective judgment is to make a system of nature, but this is not an adequate *practical* description of judicial expedience because it is not sufficient to ground the operation of reflective judgment in the totality of the system depicted in figure 4. The fundamental principle that underlies all operations of the mental processes in figure 4 is the principle that the development of intelligence and thought happens through the aforementioned process of equilibration through assimilation and adaptation. From this *practical* standpoint[5], establishing a state of equilibrium by means of the process of equilibration is a "pure purpose" *a priori* of the system. Therefore, a representation is expedient if and only if it serves the general process of equilibration, and all judgmental representations of the process of reflective judgment are generically judgments of expedience.

As a non-cognitive judicial capacity, the inputs to the process of reflective judgment are affective perceptions and its outputs are regulatory signals. There are two distinct but cooperative uses that affective perceptions can be put to by the process of reflective judgment, as we discussed above: as energetics of actions and as evaluations. An action, whether an act of reasoning (mental action) or a motor act, serves the process of equilibration by changing the condition of the system from a condition of disequilibrium toward a condition of equilibrium. Establishment of a "direction" for an action, which Kant called an act of orientation, is an act of evaluation. Stimulation of the action (and also suppression of other actions) is an employment of the data of affective perception as an energetic.

With regard to the energetics aspect of affective perceptions, the act of reflective judgment is *practical*, and as such co-involves the process of practical reason in figure 4. Piaget has shown that cognizance and the formation of that interconnected manifold of cognitions and action schemes we call a *nexus* of *meanings* obeys a "logic of actions" [53], and that cognizance begins "at the periphery," i.e. in the observation of actions and outcomes, and proceeds "toward the center," i.e. to the making of cognitive inferences [52]. The hypothesis that our mental organization contains an innate formal system of logic, e.g. either a formal mathematical or a predicate logic, is demonstrably false through experimental psychology [59]. The only "innate" logic that can be demonstrated through actual experimentation is a "logic of actions" that provides merely the foundations for the later development of logico-mathematical thinking.

This means that an "intelligent" neurocomputer *system*, in the context of human intelligence, requires a subsystem whereby it can interact with its environment not merely as a patient but also as an agent. It is not enough for it to merely be affected by this environment, but must also possess the capacity to be able to affect this environment through its own actions. The subsystem that enables the system as a whole to be in full commerce with its environment can be called the system's sensorimotor subsystem, i.e. its "body." This need for full reciprocity between the system and its environment runs contrary to the traditional paradigm of artificial intelligence, but it is an aspect that in recent years has been gaining support. It is a consideration in Woods' model of a reasoning agent, and has likewise been recognized by other researchers, e.g. Haugeland [18] and Brooks [81]. In the words of Damasio [78], the mind is embodied, not merely embrained.

William James [50] and Damasio [79] have described similar models for the role of affectivity and motoregulatory expression in the development of objective cognition. In addition to the data of the external senses,

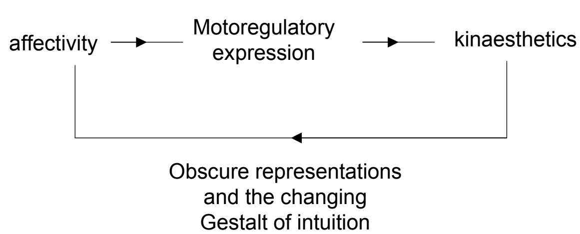


Figure 6: Affective sensorimotor loop model. Affectivity gives rise to motoregulatory expression of body states. The ensuing bodily actions produce kinaesthetic sensations that feed back through sensibility to alter both the representation of an intuition and the representation of the affective state of the system.

[5] Kant's theory is built upon a three-fold system of perspectives involving three distinct but interrelated philosophical "standpoints." These are: the theoretical standpoint, the practical standpoint, and the judicial standpoint. S. Palmquist [80] was the first to point this out, and to point out the importance of understanding this system of perspectives if one is to properly understand Kant's theory.

changes in body state (motoregulatory expression) produce sensible feedback (kinaesthetic representations) that affects both the system's affective perceptions and its objective representation of intuition. This is depicted in figure 6. Damasio goes beyond James' theory by additionally proposing that the brain contains "as if body" feedback loops, stimulated by convergence zones, which bypass physical motoregulatory expression and reproduce images of previously experienced kinaesthetic representations in sensibility. Interaction in the affective sensorimotor loop model is the reflective counterpart of regulation by means of cognitive interactions proposed by Piaget [62], which will be discussed later. Thus, one function of the process of reflective judgment is to assert connections to the system's motor faculty, and the assertion of these connections is conditioned by affective perception. The possibility of this, prior to the construction of object representations, requires particular innate "reflexes" (*a priori* sensorimotor rule schemes), and the further development of cognition under the logic of actions hypothesis requires that these rule schemes be subject to accommodation leading to assimilation of perceptions in various sensorimotor schemes.

Expedience in evaluation is two-fold, having both a practical aspect (practical standpoint) and a theoretical aspect (theoretical standpoint). The practical aspect involves the outer feedback loops in figure 4, i.e.

sensibility → reflective judgment → reason → determining judgment → sensibility
↓→→→→→→→→↓→→ motoregulatory expression →↑ .

This loop functions as an action loop, either motor action, "mental" action, or both. Upon the perception of a dis-equilibrium the system takes action (under command of the process of reason) to restore or re-acquire a state of equilibrium. However, this restoration will in most circumstances not take place immediately and a series of intermediate conditions will ensue, either approaching equilibrium for successful actions or diverging farther from equilibrium for unsuccessful actions. Inferences of practical expedience are evaluations of the current state of and direction of change in equilibrium that serve as conditions for the process of reason by which rules of action are or are not put into effect. Here we note that an inference of practical expedience is an assertion by reflective judgment, but this assertion is a condition to the process of reason, i.e. we have a series connection of rules running between the blocks of the model. The series always begins with rules of sensibility and ends with assertions re-entering sensibility or motoregulatory expression or both.

The energetics function of reflective judgment asserts a "motor set" condition, called the "meaning implications" in figure 4, but actual motor action requires the "consent" of reason through the assertion of "go conditions," called "appetitions in sensorimotor assimilation" in figure 4. Reflective judgment makes no decisions. It makes "rulings" as to what manifold of possible actions it "believes" will serve the process of equilibration, expressing this manifold through binding codes that bias some combination of diverse networks for possible activation while biasing other networks against activation. However, the actual activation of any particular arrangement of interconnected networks is subject to a veto by the process of reason. In this sense, we can say that inferences of practical expedience are "expressions of desires and needs" expressed at possibly different intensities of "desire and need." Practical reason, by exercising its veto power on some of these "desires" and allowing others to come to actualization, "makes a choice" of what actions will be pursued. Note, however, that this does not preclude the motoregulatory expression response to the motor set assertions from changing the state of the motor drive system in such a way that "as if body state" kinaesthetic feedback to sensibility is provided even if "locomotion" (external expression of motor activation) is vetoed.

Expedience in evaluation in its theoretical aspect involves the middle loop in figure 4, i.e.

sensibility → reflective judgment → determining judgment → sensibility.

The pathway from reflective judgment to determining judgment, called the "inferences of aesthetic expedience" in figure 4, have a function similar to the function of the motor set signals. These inferences bias a connection of networks serving the determining judgment process either in favor of or against activation. Again, however, the actual activation of these networks is subject to a veto by the process of reason, represented in figure 4 by the "orientations in conceptual assimilation" pathway. The critical distinction is that the conditioning signals provided by reflective judgment in this case do not lead to motor action but rather to stimulation of convergence zones serving determining judgment. As will be explained in the section on determining judgment, the stimulation of these convergence zones leads to the formation of *concepts* (which will be defined later), and we will call the system activity undertaken in the formation of object-concepts "thinking."

Here it must be recalled that a determining judgment is the process by which a generic concept is available and particular representations are found that can be subsumed under the general concept. The process of determining judgment itself makes no general determinations and must be "given" the general representations it uses. The making of a general concept is an act that must serve a logical expedience for equilibration, and the judgment of such expedience belongs to reflective judgment. Thus, the inferences of aesthetic expedience can be called "inferences of judgment." Three modes for these inferences can be distinguished:

1. Inference of ideation - the binding of specific feature networks to form an original concept of an object;
2. inference of analogy - the binding of specific characteristics found in common among other objects to a similar object. Analogy concludes from *partial* similarity between two objects to *total* similarity, i.e., things of one genus which we know to agree in much also agree in the remainder as we know it in some of the genus but do not perceive it in others. An inference of analogy is a *specification* process that expands the properties of an object already contained in its concept to endow this object-concept with other properties obtained from concepts of other objects judged to be similar to this one;
3. inference of induction - the binding of characteristics found in many objects to form the concept of a genus. Induction is a process of generalization, i.e., "one characteristic found in many objects is characteristic of *all* objects" that can be assimilated into this new genus.

Induction expands the empirically-given from the particular to the general with regard to many objects. Analogy expands from the given properties of one object to further properties of that same object. Both of these types of inferences, however, require that object-concepts already exist, and the making of a completely new object-concept is an inference of ideation.

Note that the rules of reflective judgment by which these inferences are asserted do not directly judge objects. These rules merely judge the expedience of generalizing the activity patterns present in the representation of an intuition, i.e. the expedience of binding particular feature activities together to form a concept. These judgments are therefore not based on any "copy of reality" hypothesis, are themselves non-objective, and hence should be called "subjective" or "aesthetic" assessments of sensibility. The data of the senses provide no general laws, and it is up to reflective judgment to make a unified system of experience by "pushing" determining judgment into organizing its manifold of concepts in a manner expedient to the making of such a system. It does so by evaluation of whether or not a state of equilibrium exists in the firing activities present in sensibility. Reflective judgment does not produce the binding codes that link features in sensibility; it induces convergence zones serving determining judgment to *learn* these patterns so that they can be reproduced (which, of course, makes possible the future re-establishment of equilibrium in sensibility). The signals instigating this adaptation comprise a *marking*.

C. Determining Judgment. The function of the process of determining judgment is *to construct rules for the reproduction of intuitions*. Such a rule is called a "concept." Concepts are constructed through the formation of binding codes from convergence zones in accordance with Damasio's architecture. An intuition for which a concept has been constructed is called a *cognition*. Once constructed, concepts can participate in the formation of new intuitions by stimulation of feature networks serving the sensibility process. Following Kant's terminology, the process of re-stimulating sensibility is called the synthesis of reproductive imagination (when the action merely reproduces an intuition) or the synthesis of productive imagination (when the stimulation of sensibility by multiple concepts, possibly also combined with new sense data, leads to the formation of an entirely new intuition). The cognition of an object through concepts is called *thinking*.

A determinant judgment is an act of combination by which two or more concepts are connected in intuition to form another concept. Note that the intervention of reflective judgment is required only in order to stimulate the formation of a new concept, but that the specific details of the concept formation, i.e. the combination of particular concepts synthesized into the intuition, is adjudicated entirely by the process of determining judgment. The distinction here is somewhat subtle. To use Kant's terminology, the inference of judgment carried out by reflective judgment *makes a distinct object*; the determinant judgment carried out by the process of determining judgment *makes an object distinct*, i.e. it adds characteristics to the object-concept by using other concepts as "marks" of the root concept. For a given object-concept ("given" through reflective judgment), the process of determining judgment finds other concepts (also already presented to it by past judgments) that constitute the "particulars" for the "general" object-concept. By way of analogy with formal predicate logic, the concept of the object being made more distinct corresponds to the subject term; the concepts that act as marks characterizing the subject-concept correspond to the predicate. Having said this, however, it must be immediately noted that Kantian logic distinguishes three forms of such combinations. Subject-predicate (categorical form) is one of these. The others are condition-conditioned (a form called "hypothetical" or "if-then" in conventional logic) and concept-"members of divided concept" (disjunctive form, e.g. "*C* is *A* or *B* but not both *A* and *B* at once").

None of this should be taken to mean that the process of determining judgment is merely a mechanism of formal logic. Formal logic admits no objects; it operates on "variables." As is usually said, formal logic makes abstraction of all "contingent" or "material" content and deals strictly with logical form. Determining judgment acts on representations of sensibility, which are always empirical, and its judicial operations take into account the origin, domain, and objective validity of its constructed concepts. To put it another way, determining judgment judges *objects* not "variables." It does so through the application of a set of *a priori* meta-rules for the construction of concepts. Kant interchangeably called these meta-rules by two names: "pure notions of understanding" and "categories of understanding." The two terms are synonyms. These rules do not specify objects *per se* (they are not "innate ideas"); they specify *what is required for a representation to be objective*, *what forms of conceptual representations are objectively valid*, and *how concepts are to be employed in thinking*. Because these rules specify the conditions under which a perception can be a valid objective perception, i.e. can refer to an object, Kant called this logic a "transcendental" logic. He based the deduction and definition of these rules on the epistemological principle ("meta-meta-rule") that these pure *a priori* notions should comprise a complete set of rules *necessary for the possibility of experience*.

Determining judgment involves the inner feedback loop in figure 4, i.e.

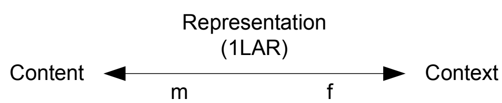
$$\text{sensibility} \to \text{determining judgment} \quad \to \text{sensibility.}$$

Its job is to form combinations of feature networks to produce the intuitions of objects (entities and events), and to do so in such a way that conscious cognitions form a unified structure. As stated earlier, consciousness is the representation that another representation is being presented. In effect this is a feature identification function that on the psychological plane could be characterized as an awareness mechanism leading to attention. Damasio defines "attention" as a "spotlighting" or signal enhancement process that generates simultaneous and multiple-site salience. As such, attention is a binding code mechanism and, logically, denotes a rule by which some feature networks receive additional excitation while others have their activities depressed (background) or inhibited (made unconscious). Because it is a rule, the attention mechanism must have a condition (an awareness condition). Here we make the hypothesis that awareness of features takes place when the feature network activity reaches a certain firing rate threshold and approaches time-lock to within some critical phase difference. Detection of this condition requires coherence detectors similar to those discussed earlier. But in the case of cognitions, there must also be a mechanism for coordination of the multiple awareness detectors. This implies the need for higher-order convergence zones dedicated to the consciousness-producing mechanism of the system and acting through the indirect feedback pathways of the network to synchronize lower level awareness detectors covering multiple regions in the feature network system. Detection of this "higher order" synchronization of convergence zones marks a condition of state in the system that can reasonably be called "apperception."

By using this neural network model schema we can appreciate the functional character of Kant's categories. In *Metaphysik Vigilantius* ([21, vol. 29, pg. 984]) Kant states, "The consciousness of the synthetic unity of the manifold in intuition in general, or (what is the same) the consciousness of the concept that contains the synthetic unity, is equal to category." Determining judgment combines particular feature networks into one overall activity schema by acting through the network of convergence zones. The resulting representation of an intuition is simultaneously accompanied by "awareness features" represented in the specific activity patterns of those convergence zones serving the apperception function. Let us call these convergence zone networks "category networks." The structure of the activity pattern in sensibility is the intuition of the object, but the structure of the activity pattern in the network of "awareness features" convergence zones serving determining judgment is the form of the determinant *judgment* of the object because this pattern is the determination of the form of apperception of the object, i.e. the system's *understanding* of the object. In the first edition of *Critique of Pure Reason* ([21, vol. 4, pg. 88]) Kant wrote, "The unity of apperception with regard to the synthesis of imagination is understanding . . . Consequently, pure *a priori* notions that contain the necessary unity of the pure synthesis of imagination with respect to all possible appearances are in understanding. These . . . are the categories."

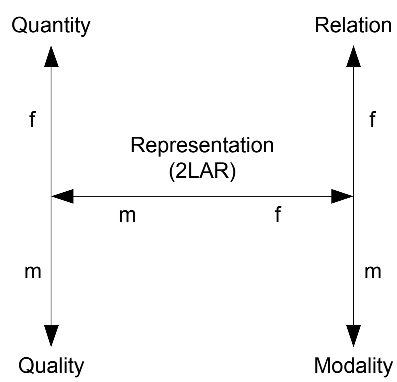
In less philosophical terms, the categories are meta-rules that determine how the system can *use* its intuitive representations of objects. They are, in this sense, representations of knowledge *a priori*, and this term means the system's repertoire of "know-how" that is necessary for the possibility of empirical knowledge of experience. Determining judgment receives orientation signals from the process of reason, and we can regard the representations in the activity patterns of category networks as state variables of the process of determining judgment. The combination of category representations as state variables, data inputs from sensibility in the imaginative re-cognition pathway and orientation signals from reason determine the operation of the process of determining judgment.

Now, by calling the categories "know-how" built into the system, the issue of knowledge representation is raised. This is a problem of well-known importance in cognitive systems work. Underlying it, however, is a more basic issue, namely, what in general does "representation" mean? "Representation" is a primitive term because, as Kant pointed out, the only way to explain "representation" is to *make* a representation. However, while circular, this issue is not viciously circular because we can obtain an operational understanding of "representation" by describing its necessary characteristics. At the first level of explanation, a representation must contain two characteristics. It must represent "something"; this is called the *content* of the representation. It must also provide a *context* for that which is represented. A representation represents an existent, and, in the first level analytical representation of representation (1LAR) just described, the content pertains to existence in the connotation of a *Dasein* (existence in the sense of something that exists) while the context pertains to existence in the connotation of its *Existenz* (way or manner in which it exists). Kant called the content of a representation its *matter* and the context of a representation its *form*. The diagram



illustrates these ideas. In this diagram, "m" denotes matter and "f" denotes form.

In analyzing representation we need not stop at the 1LAR level. Both "content" and "context" can be represented by a similar division into matter and form to obtain a second level analytical representation (2LAR). The diagram



illustrates this idea. The root representation is now depicted as having a form-of-the-matter (which Kant called Quantity), a matter-of-the-matter (Quality), a form-of-the-form (Relation), and a matter-of-the-form (Modality)[6]. Obviously we could continue this division process indefinitely, but if at some point we stop dividing and want to make use of this picture, the last legs of the division must be supplied with *rules of synthesis* that tell us how to employ this structure. Kant explained that under each terminating "title" (e.g. Quantity) there must be three such rules, which he called *momenta*[7]. The reason there must be three is because synthesis in logic always requires three terms, e.g. (1) a conditioned, (2) a condition, and (3) that which arises from the union of the conditioned with its condition. The synthesizing rules tell us how to use the representation.

A synthetic judgment combines two objects in some manner, and so for combination in judgment the content of the combination is called its *composition*, and the context is called its *connection*. Kant deduced his categories by regarding them as the synthesizing functions at the second analytical level of representation of the idea of judgment. Thus, there are twelve categories (three for each of the four titles). They are:

Quantity: unity ("one-ness"), plurality ("many-ness"), and totality ("all-ness");
Quality: reality, negation, limitation;
Relation: inherence and subsistence, causality and dependence, community;

---

[6] Kant never explicitly laid out the diagrams shown above, but this organization is implicit in all the different contexts in which he used the terms Quantity, Quality, Relation, and Modality as technical terms.
[7] One transferred use of the Latin word *momentum* is "mental impulse."

Modality: possibility-impossibility; actuality-nonbeing; necessity-contingency.

Because all four "ends" of the 2LAR diagram are needed to make *one* representation (at this level), every determinant judgment involves the application of four categories (one from each title)[8].

To the consternation of philosophers ever since, Kant did not provide definitions of these terms in *Critique of Pure Reason*, although he discussed them at length in his metaphysics lectures. For the purposes of this paper, we need not trouble ourselves with all the philosophical and psychological ramifications of Kant's categories [71]. It is enough that we obtain a functional explanation of how they instruct the system in the use of its cognitions and what consequences they hold for the subsequent actions taken by the system. In the descriptions that follow, it is important to bear in mind that every determinant judgment produces a new concept but that the categories are not the concept. The concept is a rule, the categories are *a priori* rules about rules. The concept subsists in the combination (through binding codes) of the features signaled by multi-regional feature networks. The categories subsist in the structure of concepts and state of their category networks.

With regard to Quantity: The category of **unity** means that the system will display syncretism in how it acts on the object and thinks with that object's concept. The unity of a concept is its structure. **Plurality** means the system will differentiate between objects and juxtapose of them in its thinking. Plurality produces sub-systems within a concept structure. **Totality** means the system will act and think in a manner that integrates objects, taking concepts of juxtaposed objects and combining them with each other. Totality is the coordination of objective sub-systems within one object structure.

With regard to Quality: **Reality** means the system will treat the representation in intuition as an existent in nature, i.e. it affirms an object for the concept. Kant pointed out that the condition for a judgment of reality is that the concept of the object must involve actual data of sensation, or be combined through other judgments to a concept that contained this, and not be rooted in a purely imaginary representation with no actual sensation in its composition. Reality in a concept makes that concept a scheme for assimilation of intuitions. The category of reality sets up the concept structure of an object in which intuitions may coalesce as characteristic marks even when the sensibility of the intuition is an imaginative production lacking presentment of current sensuous data. **Negation** means the judgment accommodates object-structure through inhibition of a characteristic, e.g. "day is-not dark." **Limitation** is affirmation through denial. It is a judgment of what other objects the object of the judgment is not, e.g. "Sam is not-a-German." The category of limitation assimilates opposition to the concept of a second object into the first's object-structure, whereas negation denies *characteristics* to the object.

With regard to Relation: **Inherence and subsistence** means the system will recognize different intuitions as intuitions of different appearances of the same object. It establishes connections of many concepts (concepts of the object as predicates) to a *root* concept (concept of the object as subject). **Causality and dependence** means that from the appearance in intuition of one object the system will suppose the existence of another object as that which is the condition of the appearance of the first object. The category of causality and dependence establishes time order in the reproduction of a series of intuitions. Whereas inherence and subsistence is the category of entities, causality and dependence is the category of events. **Community** means that the system will recognize two or more objects within the same intuition and coordinate their concepts through reciprocal determinations (e.g. *A* is to the left of *B* <u>and</u> *B* is to the right of *A*). Concepts of community are concepts of reciprocal relationships of coexistence at the same moment in time.

With regard to Modality: **Possibility-impossibility** means the system will regard the object of the concept as being actualizable or not actualizable in experience. **Actuality-nonbeing** means the system regards or does not regard the object as an object of actual experience. **Necessity-contingency** means the system will regard the object as either having to or not having to exist in connection with other objects in the manner represented in the concept. The *momentum* of necessity makes the outcome of the judgment of which it is part *a deductive model*. Note that the *momenta* of Modality do not add anything to the concept of the object that results from a determinant judgment. They can be better described as judgments of judgments. Possibility-impossibility promotes accommodation of cognitive structures. Actuality-nonbeing has an assertoric character that promotes cognitive equilibrium. Necessity promotes assimilation while contingency promotes coordination in applying the concept to other preparatory determinations of objects.

Despite the expression of the categories of modality in pairs of opposites, there are not six categories here but only three. The categories of Relation and Modality, as form and matter of connection in judgment, are categories

[8] For the idea of representation in general, the twelve *momenta* are as follows. Quantity: identification, differentiation, integration. Quality: agreement, opposition, subcontrarity. Relation: internal relation, external relation, transitive relation. Modality: the determinable, the determination, the determining factor.

of general correlation. In the case of the categories of Relation, the correlates are "input" concepts (i.e. the concepts acting as conditions) being combined in judgment, and one concept fills one role (e.g. the notion of inherence) while the other fills the second (e.g. notion of subsistence). In the case of Modality, the connection is not from input concept to input concept but rather from "output" concept (assertion) to the way in which this concept affects the system. Specifically, a category of Modality establishes *conditions* on forms of equilibrium for reflective judgment and reason by restricting cognitive assimilation and accommodation. A category of Modality expresses *both* what types of connections among concepts are expedient for cognitive equilibrium and what types are not in the dynamic interplay between determining judgment and the synthesis of imagination. Any later-occurring intuition that contradicts the Modality of a judgment produces a cognitive dissonance in the activity patterns in sensibility which triggers an effort to reacquire equilibrium by accommodation of the structure of cognitions. For example, an action judged as possible that fails to be achieved when tried sets up a contradiction with the previous judgment of possibility and triggers a disturbance of the state of equilibrium in the synthesis of apprehension. Categories are not thoughts, they are not ideas, and they are not concepts. They are rules governing the construction of rules for the reproduction of intuitions and rules of transformations involving these intuitions. Categories produce and maintain cognitive structures and that is all they do.

Mathematically, there are 81 possible combinations of categories that can occur in any particular determinant judgment. This does not mean that the number of possible *concepts* is 81. Concepts are as individual and as diverse as the realizable activity patterns in sensibility. The categories permit 81 different *classes* of judgments.

D. Reason. This paper uses the term "reason" to mean the master process that regulates the spontaneity of the system insofar as this deals with non-autonomic functions. The principle of this regulation is equilibration, and the process of reason constitutes the fundamental control unit of the system. As an "equilibrium engine," reason is not concerned with the representation of objects nor with specific motor acts except as these are by-products of its primary function of *equilibrating "mental" structures* that extend the physical organization of the system.

Reason accomplishes this through two distinguishable but nonetheless inseparable and coordinated processes: practical reason and speculative reason. Practical reason is wholly concerned with the determination of which acts the system will undertake. It does not determine the actions themselves. That task belongs to reflective judgment, insofar as motoregulatory expression is concerned, and to determining judgment insofar as cognitive acts are concerned. What practical reason does is to choose from the alternatives presented as possible acts on the basis of the perceived relative expedience of the *form* of these acts in serving the attainment of equilibrium. Insofar as the synthesis of cognitions is concerned, reflective judgment presents possible "matters of attention" and actions associated with them, and practical reason selects "what will be attended to." In psychological terms we can view reflective judgment as presenting "desires and needs" while practical reason exercises the prerogative of "choice."

Speculative reason regulates the employment of the process of determining judgment. Again, reason in this role is not concerned with specific objects of cognition but merely with the form of the cognitive structure. It regulates for an ideal or "standard gauge" in the determination of which concepts will be used in the synthesis of imagination to produce new cognitions. Reason acts by regulating the construction of cognitive structures according to a schema of increasing levels of better and more stable schemes of cognitive equilibrium. However, the acts of speculative reason in regulating cognitive actions are not separable from the acts of practical reason acting through the employment of the system's capacity for motoregulatory expression. Studies of the development of human cognizance clearly show that the development of understanding proceeds from the starting point of a "logic of actions" from which are obtained the root "meanings" of objects of cognition [52]-[53]. The development of cognitive structures merely extends this innate logic of actions, leading to the development of logical-mathematical operations [69], [82].

D.1 Interaction structures. The basis in science for this model comes from developmental psychology and, in particular, from the work of Piaget and his collaborators. Studies demonstrate that the "sensorimotor intelligence" of children below the age of two years is intelligence at an entirely practical level [70], [63]. There is at first no conceptual distinction made between the observables stemming from the infant's action schemes and the observables tied to the object acted upon, i.e. no distinction is drawn between ends and means. The infant exercises his or her innate reflex schemes "for their own sake" and repeats these actions in cycles, improving his or her skills through playful practice. Let the observables perceived at this stage be denoted by Obs. OS. We can diagram the process of equilibration at this stage as

$$\text{Innate schemes} \leftrightarrow \text{Obs. OS}$$

where the double-headed arrow denotes equilibration.

As various reflex schemes are repeated they form acquired habits. In time the application of two or more schemes to the same object, or objects perceived as similar, leads to a coordination of schemes

$$\text{Coord. schemes} \leftrightarrow \text{Obs. OS} \quad .$$

The object acts as the point of intersection of different schemes, and this in time leads to the formation of an elementary *interaction* where the observables in the scheme and those of the object are equilibrated in what Piaget called a Type I interaction structure [62]. What was at first the cognition of Obs. OS becomes differentiated (category of plurality) into the observables of the scheme, Obs. S, and the observables of the object serving as an aliment of the scheme, Obs. O. Note, however, that the condition for this to take place is the perception of a difference between the appearance of an expected Obs. OS and the actual Obs. OS experienced. In other words, the action must be sufficiently *frustrated* (not turn out as expected – categories of negation and causality-dependence) to cause a disturbance triggering an adaptation to restore equilibrium, but not be frustrated to the extent that a rupture occurs in the cycle of adaptation.

Two varieties of Type I interactions can be distinguished. Type IA interactions involve merely practical schemes of motoregulatory expression, and interactions of this type precede the other, Type IB, form. The reason for this is simple. Type IB interactions involve schemes of thinking, and before discursive actions are possible concepts must first be established. Physical actions and Type IA interactions lay the groundwork for this. Piaget modeled Obs. S in a Type IA interaction as the totality of observables of motion, Ms, and observable "feelings of effort", Ps. He modeled Obs. O somewhat more abstractly in terms of an observable "resistance" of the object to being assimilated into the scheme, Ro, and the observable motion of the object, Mo. Observable Ro must be regarded as the observation of a deviation from expectations, while Mo is the actual motion observed. Observable Ps is measured based on perceptions of Ro, i.e. on the degree of accommodation required to achieve the expected result.

Obs. S in a Type IB interaction is modeled as consisting of the observable, As, of the scheme of activity (e.g. seriation, classification, etc.) and observable, Fs, of the expected form or context imposed on the object by this activity (i.e. an anticipation). Obs. O in this case is also modeled with a resistance Ro, but instead of being merely a motion the object has an observable modification, Mo, that includes perceiving the form of connection.

Figure 7 illustrates Piaget's model of the Type I interactions. Functions *a* and *b* are observable relationships between Obs. S and Obs. O. Function *a* can be called the awareness observable and it denotes that the object's resistance to assimilation leads to attention being paid to and a cognition formed of the scheme. Function *b* can be called the anticipatory function or "expectation" observable and designates the perception of accord or discord according to whether or not the object matches up with the "desired outcome" intended for the scheme. Function *a* is in the role of an "energizing function" while *b* is in the role of an "evaluation function."

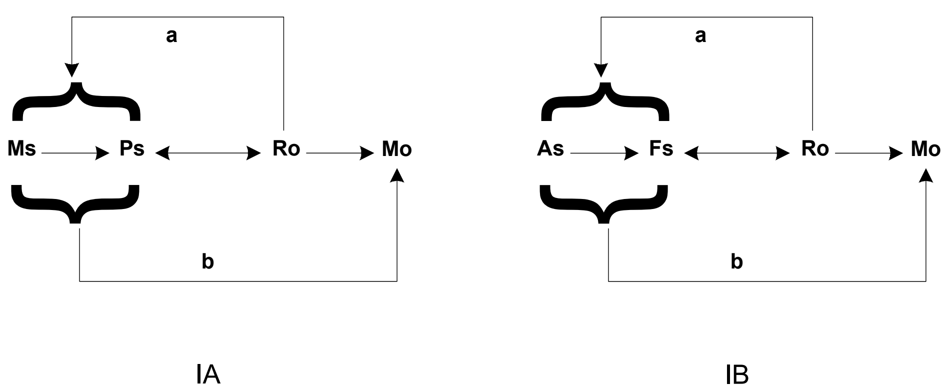


Figure 7: Type I interactions. IA: Interaction in physical action. Ms = movement observable. Ps = feeling of effort observable. Ro = perception of the resistance of the object. Mo = movement of the object observable. Function a = awareness observable. Function b = anticipation observable. Obs. S = Ms→Ps is observed as a single complex, and the effort Ps is measured as a function of the perceived resistance Ro. IB: Interaction in mental action. As = activity scheme observable. Fs = form imposed on objects by the subject. Ro = resistance of the object to being placed in the desired form. Mo = modification of the object. Type IB interactions include mental operations of thinking. Obs. S =As→Fs. The double-headed arrows denote the process of equilibration. In both types, Obs. O=Ro→Mo.

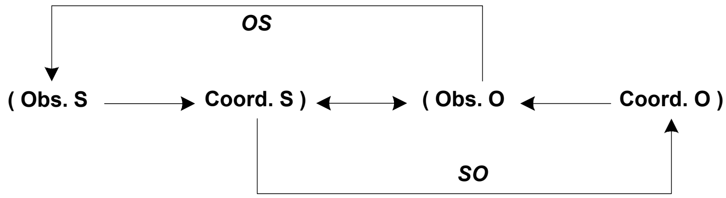


Figure 8: Type II interactions. Obs. S = action observable = Ms→Ps or As→Fs from type I interactions. Coord. S = inferential coordination of actions. Obs. O = object observable = Ro→Mo from type I interactions. Coord. O = inferential coordinations of appearances of object. Process OS = scheme of awareness; it contains interactions of Type I. Process SO = causality coordination scheme; it contains interactions of Type I. Process OS and process SO both run in each direction and the arrows merely indicate the dominant direction. The double-headed arrow denotes the process of equilibration.

All the terms in figure 7 represent observables, i.e. direct "readings" the system makes on the basis of sensorimotor perceptions of both the objective and affective types. Successful assimilation brings Obs. S and Obs. O into an equilibrium. However, this equilibrium is not very robust and subsequent activities will often turn up additional disturbances that lead to *inferences* of coordinating relationships for both the scheme (Coord. S) and the object (Coord. O) [62]. When these inferences are added to the Type I structure we obtain a Type II interaction. This is illustrated in figure 8. This structure is more robust than the Type I interaction and is enriched by the inclusion of multiple Type I scheme structures in awareness and coordination processes (process OS and process SO in the figure). These processes are essentially schemes of observation and concepts of cause and effect.

Type II structures become increasingly more general and better equilibrated with additional experience. Piaget found there is a hierarchy of increasing and better equilibriums as the structures undergo development. This is illustrated in figure 9. There are two regulatory factors at work in this elaboration of interaction structures. The first is comprised of what Piaget called compensating behaviors [62]. The second consists of what might best be called schemata of reasoning because they define the standard for what constitutes "better" equilibration in cognitive structure. They are regulative principles that functionally impose what Kant called a "transcendental Idea (*Idee*) of pure reason." However, these Ideas are not to be regarded as concepts of objects. They are better described as orientations for the direction to be taken in the employment of determining judgment for the purpose of improving the general structure of an interconnected manifold of concepts through the process of assimilation.

D.2: Compensating behaviors and psychological constitutive functions. Compensating behaviors are based on four constitutive psychological functions and correspond to three species of assimilation. A psychological function is defined as the expression of a scheme of assimilation of actions. Functions express links that are proper to schemes of action and essentially express dependences [61]. Piaget identified four innate constitutive psychological functions denoted B (association coordinator), W (repetition coordinator), I (identification coordinator), and C (substitution coordinator).

Any action that modifies an object *x* into some *x'* forms an ordered pair (*x*, *x'*), and this is the most basic form of constituted (made) function. Coordinator B expresses this adaptation. Note that the accommodation *x'* becomes a subsystem of structure *x* and that *x* does not lose any of its previous structural properties. In logical terminology *x'* is a species of *x*. Coordinator B is involved in the formation of every constituted function.

Coordinator W forms *functional* assimilations through repetition of an action, i.e. "practice makes perfect." Piaget found that at the level of sensorimotor intelligence infants exhibit "circular reactions" of increasingly complex behavioral forms, and adaptation works by means of these repetitions of an action. Exercising a scheme consolidates it and brings it into equilibrium. The basic interactions discussed above are formed through this process. In this way the system supplies its own training data.

Coordinator I forms *recognitory* assimilations. It is the constitutive function at work in the synthesis of imagination in re-cognition, but it also applies to the identification of an action scheme of motoregulatory expression in response to presentations in sensibility. Before the development of language an infant exhibits recognition of an object by expressing an action scheme to which that object is joined, but without necessarily trying to apply that action scheme to the object [70]. Two points are worth bearing in mind here. First, meanings are first formed and applied to objects through a "logic of actions." Coordinator I is the constitutive function upon which such meanings are formed in reflective judgment by joining the perception of an intuition, *i*, to a scheme of motor expression, *m*. The resulting function can be denoted by I(*i*, *m*).

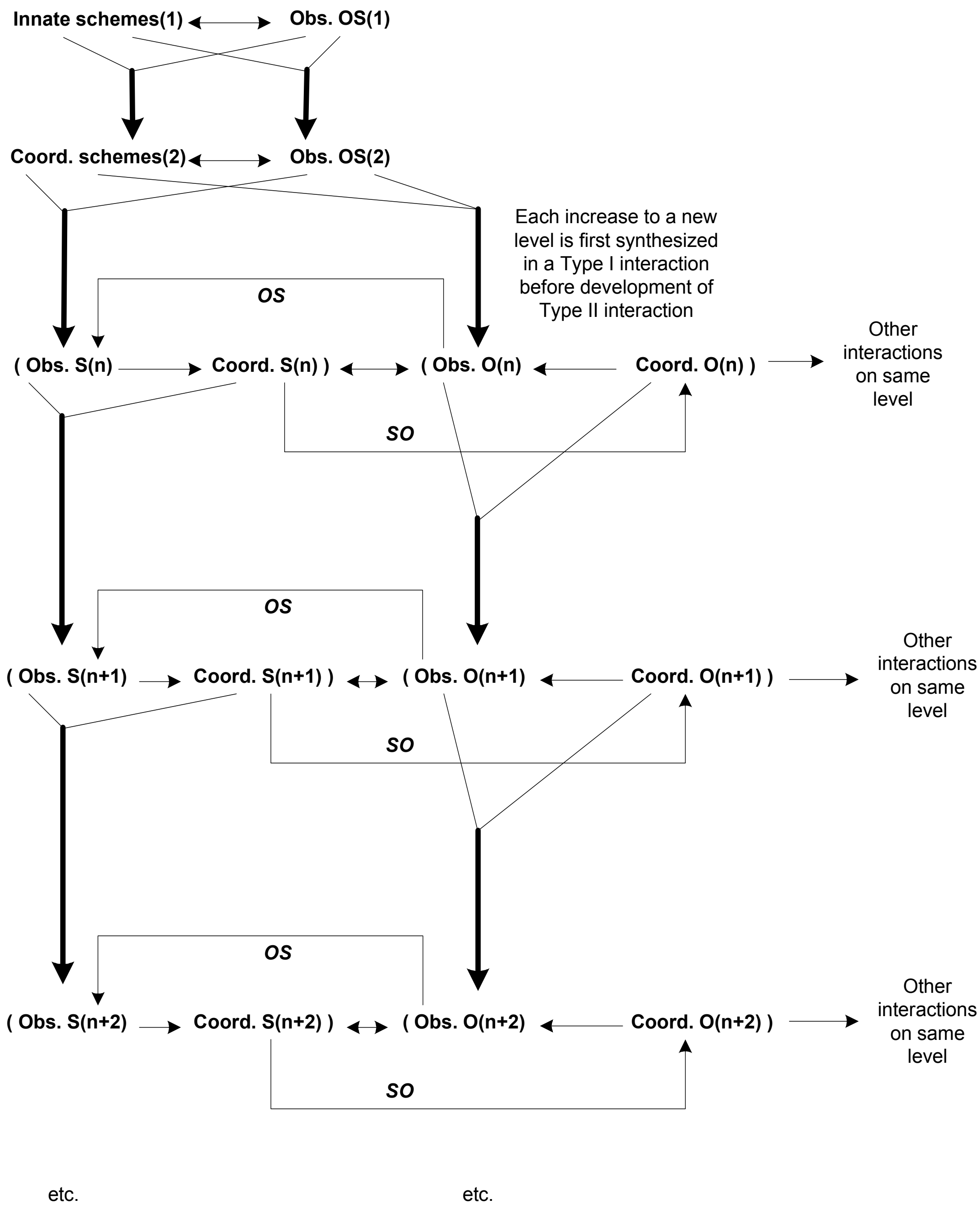


Figure 9: Hierarchy of increasing equilibrations. Innate sensorimotor schemes are initially uncoordinated and the observable Obs. OS does not distinguish between the action and the object. The object provides a common point of intersection by which schemes become coordinated. At this stage there is still no action-object distinction. Increasing equilibration going to the next stage takes place first in the form of Type I interactions. The Type I interactions are then elaborated into the Type II interactions shown here. Only after a Type II structure has been formed is it possible to increase to the next higher level of equilibration. Different interaction structures at the same level of equilibration are interconnected through inferences of object coordinations. The process generates subsystems within an equilibrated structure, but adaptation always preserves previous assimilations in this structure. Double-headed arrows denote the process of equilibration.

Coordinator C forms *generalizing* assimilations. This form of assimilation takes place by means of the substitution of a new object in place of a previous object assimilated into the scheme. Piaget identified a hierarchy of forms possible with coordinator C. In simple substitution an object x is selected in place of another object y in the application of the scheme. In reciprocal substitution x can be substituted for y and vice versa. In inversal substitutions more complex structures can be substituted, e.g. xy can be substituted for yx and vice versa.

The forms of assimilation supported by these coordinators correspond to three types of equilibrations: 1) equilibration between assimilation in schemes of action and accommodation of these schemes to objects; 2) equilibrations resulting from interactions between subsystems; and 3) progressive equilibrations between differentiations (which produce subsystems) and integrations (which unite subsystems into the totality of the overall structure). (2) differs from (3) in that (2) deals with collateral interactions at the same level while (3) deals with the hierarchical progression of levels of equilibration. Now, equilibrations lead to equilibrium and this presumes that prior to acts of equilibration there must have existed some previous form of equilibrium and some disturbance necessitating an equilibrating response. The reaction to a disturbance is called a "compensation," and Piaget identified three types of compensating behaviors.

Compensation type α is the simplest form of compensation behavior. It consists of simply canceling the disturbance, either through some movement opposite to the disturbance (in a physical scheme) or by simply *ignoring* or *avoiding* the disturbance altogether. If we let d denote the disturbance then compensation α corresponds to $-d$ (negation of the disturbance) so that the action taken is denoted by $d - d = 0$. Type α behavior concerns only assimilation and accommodation within a single scheme. Type I interactions are the result of type α compensations.

Compensation type β is behavior that integrates the disturbing element into the system. Here there is an accommodation of the scheme and what was a disturbance now becomes a variation. Type α compensation can be regarded as a form of grouping by ignoring differences, e.g.

$$A + A = A \, ; \qquad A - A = 0.$$

Piaget calls structures of this type "algebraic" in the sense that a grouping takes place here. Note, however, that the "algebraic" operation from this type of compensation does not possess the associative property of arithmetic, i.e.

$$(A + A) - A = A - A = 0; \qquad A + (A - A) = A + 0 = A.$$

Type β behavior, on the other hand, produces an *order structure* by establishing reciprocities that transform contradictions into mere contraries. A relationship < based on recognition of variations, i.e. $A < B$, also produces a reciprocal relationship > such that $A < B \Leftrightarrow B > A$. The simplest form of this, stated in the language of formal logic, is the disjunction (A and not-B) OR (not-A and B). For the structure of concepts this calls upon the category of community, and in terms of actions this structure represents the matter of a choice. The reversibility inherent in a type β compensation is neither inversion nor negation. Nothing is negated here. Rather, it is an ordering which constitutes a primitive form of seriation.

Type γ compensation is a "superior" form of compensation in which variations can be anticipated without having to be actually confronted in fact. To lead up to type γ compensation, a disturbance must first be transformed into a variation (order structure), and then possible variations in order structures (application of β to the order structures) can be classified (type α compensation). This gives rise to the possibility of cancellation of variations. For example, let T be an operation performed on object A. Let the result be denoted TA. Let T^{-1} be the operation that cancels the effect of T. Then we can obtain the structure $TA \Rightarrow T^{-1}TA = A$. In type γ compensation we find both: 1) a differentiation of substructures (accommodation) and 2) integration of substructures in a superior structure (assimilation).

D.3: Orientation of judgment by transcendental Ideas. Next we turn to the regulatory principles for the employment of determining judgment. In the discursive formation of concepts in the inner loop of figure 4, the process of determining judgment must stimulate the re-production of intuitions whose feature fragments can thereby participate in the formation of the new intuition that will become the new concept. However, determining judgment as a process subsumes particulars under a given general. This raises the issue of how determining judgment can come to re-present intuitions expedient for the construction of the system's world model. As Kant put it, determining judgment cannot determine its own employment. This task belongs to the process of speculative reason.

Now, reflective judgment judges expedience in sensibility, but this judgment must be based on some form of regulatory rules *a priori* (because such rules are necessary for the possibility of experience). Furthermore, these regulatory rules must be such that their regulations lead to higher and better levels of equilibration in cognitive structure. Kant called these rules the transcendental Ideas and stated them as general regulatory principles of pure reason. Taken together they describe for us the "structuring theme" of what a perfect equilibrium in the manifold of concepts aims to produce. The form of such a manifold is that of what Kant called a "transcendental Ideal." There are three general regulatory principles. They are:

1) The Idea of unconditioned unity of the thinking agent;
2) The Idea of unconditional (absolute) unity of the series of conditions of appearance;
3) The Idea of absolute unity of the condition of all objects of thinking in general.

The first Idea says in effect that all understanding involves the reference of the object to the conditions of the thinking subject (i.e. the intelligent agent). We recall that reflective judgment judges expedience in terms of whether or not, and in what way, the affective perception of the state of the system serves the process of equilibration. The conditions concern not only the perception of the state of the system but also what the system is inclined to do about it (i.e. connections to possible motoregulatory expressions). In terms of the four heads of the 2LAR structure, these conditions are: continuity in apprehension; compatibility of representations; anticipations of outcomes; and symbolic meanings vested in intuitions (originally founded upon a logic of actions in motoregulatory expression).

The second Idea describes the ideal form that combinations in the manifold of concepts move toward. *A parte posteriori* the combination of concepts in a manifold proceeds from the concept of a condition to the concept of a conditioned. This direction is within the capability of determining judgment to determine. But *a parte priori* the combination runs from the conditioned to its general conditions, which requires a generalizing judgment belonging to the process of reflective judgment. What determining judgment must do is provide the matter in sensibility for such a judgment. It does so by stimulating the re-production of sensible intuitions (i.e. by reactivating feature networks connected by binding codes) through reproductive imagination. The concepts that are able to serve for establishing a unified series of conditions *a parte priori* are those which contain the possibility of such a construction. We can picture such a series of combinations in terms of a graph in which the nodes represent particular concepts and the arcs denote the determinant judgments of combination. Nodes on the same level that combine immediately to the concept of an object are called the object's coordinate concepts or "marks." They produce "width" in the concept of the object. Those which make a remote connection to the concept of the object through a series of intermediate concepts are called its remote concepts or remote marks. They produce "depth" in the clarity of understanding the object. A concept is said to "contain" all the concepts that connect to it in ascending the graph. The "sphere" of a concept is the collection of all concepts it contains. The marks of a concept contained in its sphere are said to be the concepts by which the system "understands" the object of the concept. Root concepts of an object occupy the lowest level in the series and "contain" all the concepts that connect to them *a parte priori* (i.e. paths coming into the concept; connections *a parte priori* are said to "ascend" the graph). Root concepts are tie points for all concepts "understanding" them and they represent empirical objects. This is illustrated in figure 10. The second transcendental Idea regulates for finding conditions of an object (remote marks in the graph) that are themselves *unconditioned* by any concepts *a parte priori*. This is of course a never-ending process, but the system does not know this *a priori* and *presumes* it has an absolutely complete series of concepts so long as the equilibrium of its structure is not disturbed. The second Idea regulates combinations of contexts for the object.

The third Idea in effect regulates for the establishment of an absolutely complete system of *meanings* for objects. Empirically speaking, the meaning of an object is "what can be done with it, what can be thought with it," and so on. In other words, all meanings are at root *practical* and a concept is said to provide meaning when it is tied (through reflective judgment and reason) to schemes of action. For a baby, a rattle is something to suck, something to shake, something to throw, etc. For an adult, a "number" is something to count with, something to produce other numbers with (i.e. do arithmetic with), etc. Actions need not be and often are not "physical" actions in the sense of producing observable locomotion. A great deal of education is devoted to helping students establish maxims of reliable *thinking*, and such a maxim is nothing else than a *scheme* of action for thinking about something in a particular context. What a person is able to think depends on what concept structures he has available for use in thinking. In terms of outcomes in thinking expressed according to the 2LAR structure, the third Idea mediates for: determination of objects and appearances; affirmation or denial of concepts of "what" the object is and is not; the context in reality of the object, i.e. as a thing (entity), event, or state of nature (world model); and the determination of systematic coherence in the concept of the object with respect to its contexts in nature.

Now when an intuition in sensibility is produced for which a corresponding concept already exists (from past experience), the process of determining judgment is capable on its own of summoning up all the other concepts combined with to this concept through previous judgments. Binding codes in the category networks make this possible. This is experimentally evident from the syncretism exhibited by childish thought [83]. However, some of the concepts that can be recalled into the synthesis of apprehension can conflict with other such concepts or with the data of the senses. Such conflicts constitute a disturbance upsetting or preventing equilibrium from being established in the interplay between the synthesis of imagination and determining judgment. Consequently, this disturbance must be either removed (type α compensation) or accommodated (type β compensation). Furthermore, the synthesis of apprehension is a process rather than a single step, and new concepts synthesized during this process can call up additional concepts, as e.g. in the case of a new concept of a mark that also happens to be contained (implicitly) in the intuition of a different concept. For example, concept H in figure 10 is a remote mark of both concepts A and B and so the recall of concept A can stimulate the recall of the concepts connected to concept H (both immediately and remotely). Obviously this can quickly lead to chaos in the operation of the inner

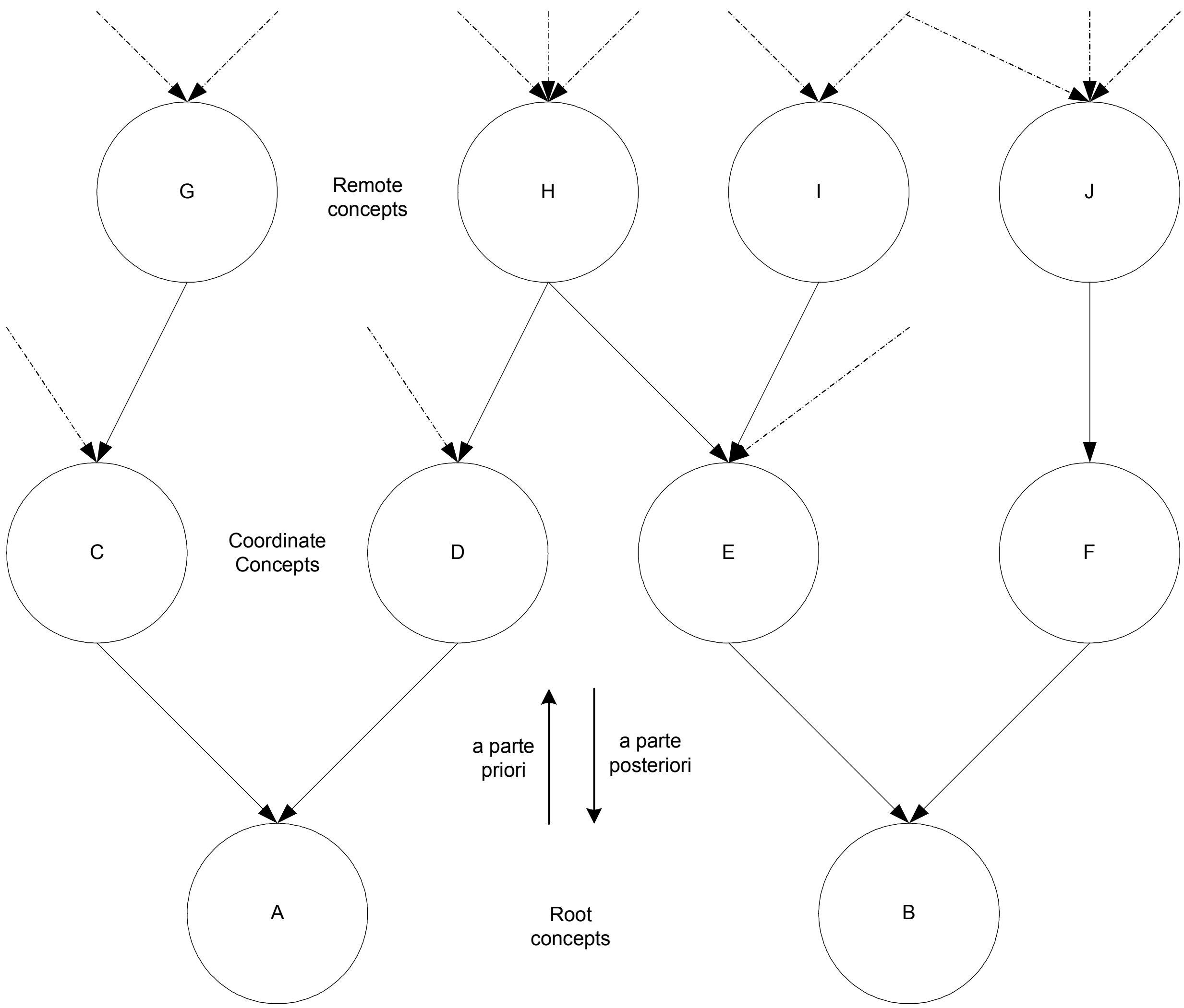


Figure 10: Graphical illustration of connection of the manifold of concepts. Nodes represent concepts. Arcs represent combinations of determinant judgments.

loop of figure 4.

The task of reason is to regulate this process and prevent the rupturing of the thinking cycle. The process of speculative reason does this by exercising a *veto* over the concepts that determining judgment is capable of recalling through reproductive imagination. In one sense this is an attention function, i.e. a coalescing of the thinking process around *one* object by preventing concepts of other objects from dominating the process of thinking. But on the other hand, the *perfection* of the structure of concepts is an aim of reason, and in this context "perfection" means the establishment of higher and better levels of equilibration (e.g. figure 9). Acting to perfect this structure is type γ compensation. What we have here is in effect a complex optimization problem with "optimum" defined by whether or not the resulting adaptation of the structure of the manifold of concepts "moves in the direction" of Kant's transcendental ideal as described in the transcendental Ideas above. The regulation of reason "strikes the balance" between Piaget's two principles of assimilation and accommodation. Note here that in this action the process of reason is concerned solely with the form of possible structures of the manifold of concepts and not at all concerned about the objects it represents. Reflective judgment *evaluates* different possible changes to the manifold, but reason *orients* the choices to be made available and uses the transcendental ideal as a standard gauge for choice.

E. Network Adaptation Considerations. The considerations given above have implications for the paradigm of adaptation of the system. Some aspects of the process of equilibration discussed above, e.g. Process OS in figure 8, are reminiscent of neural network strategies employed today, such as the use of "critics" [87]-[88]. Unfortunately the problem of making a PCNN adaptive is presently in a very primitive state. To date no widely-popular adaptation scheme comparable to the widely accepted methods such as the Widrow-Hoff rule for the generic connectionist neuron model has appeared, and no network learning scheme comparable to the various widely used schemes, such as backpropagation, for the generic connectionist neural network (GCNN) has been proposed. Such schemes as have been proposed are generally based on the Hebbian learning model or the Informax principle. Of these schemes perhaps the best known is the Bienenstock-Cooper-Monroe (BCM) rule, but BCM is an algorithm developed for and most easily applied to firing rate neuron models. It can of course be imported into a pulse-coded neuron (PCN) structure, but in doing so considerable hardware expense is incurred. It can likewise be "off loaded" to a special purpose microcontroller whose task is to manage weight adaptation in the network, but this still leaves us with the issue of building controller interfaces into each neuron, bus connections running from the neurons to the microcontroller and, probably most important of all, presents us with the problem of providing a means to store the synaptic weights in each neuron. Because the synapses usually consume the majority of chip area in a PCN, followed by interneuronal wiring, this is a key issue for adaptable PCNNs.

There are several reasons to ask whether the type of paradigm for adaptation in a PCNN does not need to be fundamentally different from the tradition that originated from the Rosenblatt/Widrow-Hoff schemes. These schemes were developed at a time when our knowledge of the dynamical character of biological synapses was quite primitive compared to today. Hebb's hypothesis was at that time merely an empirical speculation based on a few important but still very limited experimental findings. Considerable support for Hebb's hypothesis was gained with the discovery that NMDA channels in neurons do in principle seem to operate in a manner consistent with the Hebb model. However, this by no means proves that the NMDA-Hebb model applies universally in biological neurons. In point of fact, it has proven surprisingly difficult to verify the presence of this mechanism in most regions of the central nervous system.

A second consideration arises from the issue of the so-called "neural coding" scheme in biological networks. In the early days, neural network architecture was based primarily on the idea of the so-called "grandmother cell" [40]. The hypothesis was that signals upstream in the network were integrated by downstream cells whose firing activity indicated the recognition of particular features or objects. Usually this model was accompanied by the supposition that the "neural code" was a firing rate code. With the grandmother cell model it was very natural to look at adaptation "locally" as is done with the Perceptron and Widrow-Hoff rules. It was equally natural to look at network adaptation using schemes that by and large were merely means of conveying either an error signal or a desired response signal to the local neuron adaptation mechanism. Let us call this the "local adaptation" paradigm.

This local adaptation paradigm is also followed by recurrent networks, e.g. Hopfield nets, ART networks, and so on. Because most such networks employ unsupervised learning, the local adaptation paradigm became adopted for unsupervised learning as well. So it is that in both cases the local adaptation paradigm dominates neural network theory today. In addition, the conventional approach tends to not make use of any *a priori* information that might be available for the "baseline" synaptic strengths that are appropriate to the particular system. The thinking here is twofold: 1) the availability of general and robust adaptation algorithms would seem to make it relatively less important to pay attention to the initial values of the synaptic weights (unless the resulting transient response as the network "pulls in" to a nominal operating state is crucial); and 2) many engineering applications of neural networks are such that it is easier to simply train the system than it is to figure out analytically what initial weight settings might be most appropriate.

For the network architecture proposed in this paper, the signal processing scheme is much less a matter of individual neuron connection specifics and much more a matter of tight binding within relatively small subnetworks (cell assemblies) and cooperative responses among multiple larger subnetworks. This goes somewhat beyond earlier models of distributed coding schemes, where a particular subnetwork was regarded as performing vector encoding of entities and events (a "clan" system rather than a grandmother cell) [40]. This is because the biological scheme appears to use sensory subnetworks to represent only feature fragments, and the representation of entities and events subsists in the coordination of multiple feature segment networks.

In this scheme synaptic weight adaptation is clearly required of the CZ assemblies. This is because their task is to link very specific feature networks having very specific time-locked firing patterns. However, it is much less clear to what extent plastic adaptation is required within feature segment networks. Particular sensory modalities and motor excitation functions have something of a "hard link" to their input pathways (for sensory networks) or output pathways (for motor networks). It is entirely reasonable to propose that such networks do not require much in the way of adaptation capabilities. Whether or not this is so is largely dependent on how sensitive the generation of time-locked firing activity is with respect to synaptic weights within the network. Here it is worth pointing out that elementary feature-detecting and image processing PCNNs explored to date have extensive lateral and recurrent connections within the assembly, and that the function of these interconnections is mainly to enforce synchronized firing in response to a limited set of input conditions. To do so existing network models, such as the Eckhorn/Johnson network, rely on short term *elastic* modulations of *all* the data pathway synapses rather than adjusting individual synaptic weights. In artificial PCNNs reported to date this modulation takes the form of what is in effect multiplication of all the data path weights by the same multiplying factor.

This property, which was deduced from biological neural network characteristics, becomes even more interesting when it is coupled with another biological fact that was just beginning to be appreciated in the late 1970s. Biological neurons come equipped with a wealth of ion channels whose principle effect is short-term to intermediate-term signal modulation rather than synaptic weight change [30]. There are two primary biological channel types that execute this function: voltage-gated ion channels (VGCs) and metabotropic signaling channels (MSCs). VGCs tend to raise or lower the excitability of the whole neuron, produce phase delays and/or firing rate accommodations in its firing response, or produce specialized responses, e.g. ON-responders or OFF-responders. None of these mechanisms requires adjustment of synaptic weights nor variable weight storage.

MSCs are pure control pathways. They employ specialized neurotransmitters (e.g. serotonin, dopamine) and/or neuromodulators (hormones) that trigger biochemical cascade reactions that alter the response characteristics of the cell. Some of these responses can involve changes in the effective synaptic weights, but in most cases this is, again, an effect that tends to be at the "whole cell" level or possibly at the level of multiple data synapses, e.g. throughout a dendrite. Metabotropic effects tend to be slow at the onset and last for a much longer time than is the case for normal ionotropic (data pathway) signaling. Interestingly, the relatively few provable cases of long-term synaptic weight changing mechanisms caused by metabotropic signals are *presynaptic* rather than postsynaptic. This implies that the effect can be modeled at the output of the cell and does not need to be replicated individually at every synapse by means of a weight-storage mechanism (which again is beneficial in terms of the size of the artificial neuron cell and reduces or eliminates the problem of weight storage).

In addition, the past few years have revealed yet another interesting biological mechanism that might require us to significantly re-think how we view neural network adaptation. It now appears likely that *glial* cells (long thought to provide only a life support function for neurons or to improve axonal signal speed) actually participate in signal processing. Although glia do not produce action potentials, they are now known to possess as rich a variety of ion channels as do neurons, to respond to synaptic activity, and to "broadcast" that response over a very large area. For example, it has been verified that glial networks produce "calcium waves" which deliver Ca^{2+} ions to the extracellular environment of the neurons. Ca^{2+} is one of the most potent chemical effectors in biology, and changes to a cell's Ca^{2+} environment can have potentially great effects on the signal processing behavior of the neuron. Again this would be a global "whole cell" rather than local "specific synapse" mechanism and potentially another biological substrate for non-Hebbian learning.

What all these facts suggest is that perhaps what is needed is a radically different paradigm for neural network adaptation. Such a paradigm would incorporate several ideas:

1. The adaptation algorithm should be "locally global" – i.e. it should work on the principle of modulating entire cell groups rather than adapting individual cells or synapses;
2. Classical Hebbian synaptic weight adjustment mechanisms are likely to be confined to a relative few special cells located primarily or perhaps even entirely in CZ subnetworks (and not necessarily in all neurons contained in that network) and/or in specialized networks serving a non-specific function, e.g. a "hippocampal network";
3. Greater emphasis is to be placed on the elastic modulation properties of the majority of the neurons in the system rather than on long-term changes in synaptic strength;
4. The adaptation-producing mechanism for long-term changes in synaptic weights at those cells capable of such plastic changes is likely to involve specialized "control cells" that signal metabotropically;
5. The stimulation condition for the onset of plastic changes in the network should be based on coherence detectors, i.e. cells that detect the presence of time-locked, multi-region firing patterns since these patterns are the representations of entities and events;
6. Data path processing should be dependent on feedback signals from downstream CZs that act to *reproduce* firing patterns originally stimulated by sensory inputs alone.
7. A "critic" or "value" system should be incorporated, and this system functionally would be part of the process of reflective judgment.
8. An overall system of regulation (the process of reason) is required that is capable of holding the thinking activities of the system on one or a few "matters of attention", orienting the interplay of imagination and determining judgment (based on the states of category networks), and integrating discursive actions with motoregulatory expressions ("logic of actions"). Such a regulator constitutes the substrate of a psychological "volition mechanism" for the system [50].

Thus, it may well be the case that Hebbian mechanisms play only a small part in the overall landscape of PCNN adaptation. The idea of "locally global" adaptation mechanisms is a paradigm that lends itself to economies of scale in terms of neural network hardware since one structure would be able to serve to control synaptic responsiveness in many neurons.

This paradigm also has its implications for neural network design techniques. Because only a relative few neurons would actually possess plastic weight change capability, a far greater premium would be placed on the use of *a priori* knowledge of the weights of fixed synaptic connections. Here we see the potential for an interesting partnership between PCNNs and the techniques currently employed in GCNNs. The great advantage of connectionist network theory is the relatively small amount of *a priori* knowledge required of its design. These networks are in a sense "self-designing" through their adaptation algorithms. What I suggest is the following: That GCNNs can be used off-line during the initial design process to find "design center" weights for the feature segment subnetworks and for the initial settings of CZ neuron synaptic weights. These weights would then be translated into equivalent synaptic weight settings with the PCNN. Provided that the design of feature subnetworks is robust in the face of weight variations, and that modification of the system's signal processing behavior is primarily through elastic locally global mechanisms, most of these weights can be fixed, leaving only a relatively few neurons requiring plastic adaptation. Because fixed-weight synapses are easily implemented using current mirror techniques, the development of VLSI standard cell libraries can therefore be used to make the design of neurocomputers no more difficult than is the design of any other VLSI system.

V. Research Issues

The model presented in the previous sections defines in a general and often qualitative sense what the character of a general purpose neurocomputer architecture should look like. This is a description at the top level of the system, but much remains to be done before we can say that we have a design. The design of feature networks is one obvious task. To a limited degree we have some idea about these networks from work already done in PCNN applications in image processing, particularly in the work on using a PCNN for image segmentation. But this is only a small part of the total task at hand. We must figure out how to achieve stable, time-locked multi-region firing patterns that link features into object representations. This means figuring out how to implement convergence zone networks and their binding codes, as well as how to "imprint" the connections of features in convergence zones by adaptation of convergence zone neurons. In this task I make the hypothesis that various elastic *modulation* features of neurons in the network will be required as much as will be more traditional, and possibly Hebbian-based, long term adaptation of synaptic weights in particular convergence zone neurons. There is also the question of the organization of different levels of convergence zones that will be required as well as the issue of the functional and connection properties of the non-specific (and probably mostly metabotropic) networks that must play roles analogous to subcortical brain structures such as the hippocampus, hypothalamus, or limbic structures.

It has also been noted that we will require special subnetworks implementing reflective/affective functions. A few generic classes of these have been pointed out here: coherence detectors, incoherence detectors, consciousness flag networks, attention flag networks, networks mediating motoregulatory expression, etc. The detection functions called for here essentially amount to detecting frequency and phase coherence, probably within some limited range of frequencies, among signals converging at the specific convergence zone. At the present state of the science one likely mechanism for implementing some of these detection functions is the neural network counterpart of a phase locked loop (PLL). It is already known that very simple, small pulse-coded neurons (PCNs) can implement the logic functions for simple central pattern generators (CPGs) and for standard digital phase detectors (PDs). A CPG can be used to establish the reference frequency for the output firing rate of a convergence zone input neuron (CZ) in response to time-locked input signals (i.e. determine a bandpass filter function), and a PD comparing the reference CPG with the output signal of the CZ can be made to fire in response to phase delay errors between them. A schematic illustration of the basic idea for this type of detector is shown in figure 11. A bank of these detectors (all using a common CPG) acts to form a somatopic “map” of what feature network regions were doing so far as coherence is concerned. The inhibitory and excitatory outputs of each bank are fed back to multiple feature networks. The inhibitory signals tend to shut down the feature networks, and if their sum is stronger than the excitatory feedback it will. But if the excitatory output is stronger, it will tend to reinforce the firing of the feature network. In effect this is a winner-take-all scheme that works by frequency modulating the firing rate of the feature networks – in effect, using them as the counterpart of the voltage-controlled oscillator (VCO) function of a PLL. The CZ bank with the best coherence or least incoherence should gradually “win” and establish its preferred stable firing pattern. If the burst rate of the coherence detector matches that of the CPG it should be able to enforce frequency lock between the winning CZ subnetwork and the feature networks. If NONE of the converging feature network patterns are close enough to coherence, their links ALL get shut down. Such a scheme also sets up the basis for unsupervised Hebbian learning in the CZ input neuron. Once a stable feature is selected, and provided that the feature is presented for a long enough period of time, the output of the coherence detector can be used to trigger an adaptation process acting on the CZ neuron that recognized the winning pattern. This “imprints” the connection in the CZ neuron. Inputs that were actively participating get their weights increased, those that were not firing or firing at a significantly slower rate get their synaptic strengths depressed. Because the coherence condition is synchronous time-locked firing activity from the selected feature networks, the synaptic weight changes of the CZ neuron synapses need only distinguish between “winner” and “loser” pathways, thus requiring at most only *two* adjustments to the synaptic linkages (one providing potentiation of all the “winning” pathways, one depressing all the “losing” pathways). It might be sufficient to merely suppress losers.

Above all, and probably the most difficult yet crucial problem, is the problem of defining and solving the complex optimization problem that is the centerpiece characteristic of the process of reason. These issues have not yet been adequately addressed in the context of a PCNN architecture, much less in the context of an architecture based on the convergence zone model. Most of these problems are still ill-posed and require a great deal of refinement in their mathematical expression. **It seems very clear that progress toward the realization of the paradigm presented in this paper is most likely to come from small steps.** We must break down the landscape of the blocks and processes illustrated by the figures in this paper, give precise statement to the particular sub-problems that must be solved, and from these solutions attain to a more refined understanding of the problems at the next higher level. This procedure is only partially reductionist; the “top level” principles do not yet have clear and explicit mathematical statement. The general framework presented here does enable us to root out specific lower-level questions and answer them, and in this regard the research is reductionist. But the solutions to these problems will implicate the definitions of problems at the next higher level, and in this sense the research must also be constructionist.

For most of the history of neural networks the biggest obstacle to solving the optimization problem as been the NP-complete nature of perceptron network learning. Solutions that work for small-scale networks do not scale up well for larger problems. There have been many different methods employed in trying to solve large-scale problems [84], but we are still well short of what is required for systems involving hundreds of thousands of input and output variables. Our most important mathematical tool for addressing the scaling issue is the Bellman principle and optimization of the HJB equation. However, HJB leaves to our own creativity formulation of the cost function to be optimized (within some general constraint conditions). Perhaps a divide and conquer approach, generalizable to large scale problems, could provide relief from the “curse of dimensionality” issue inherent in the

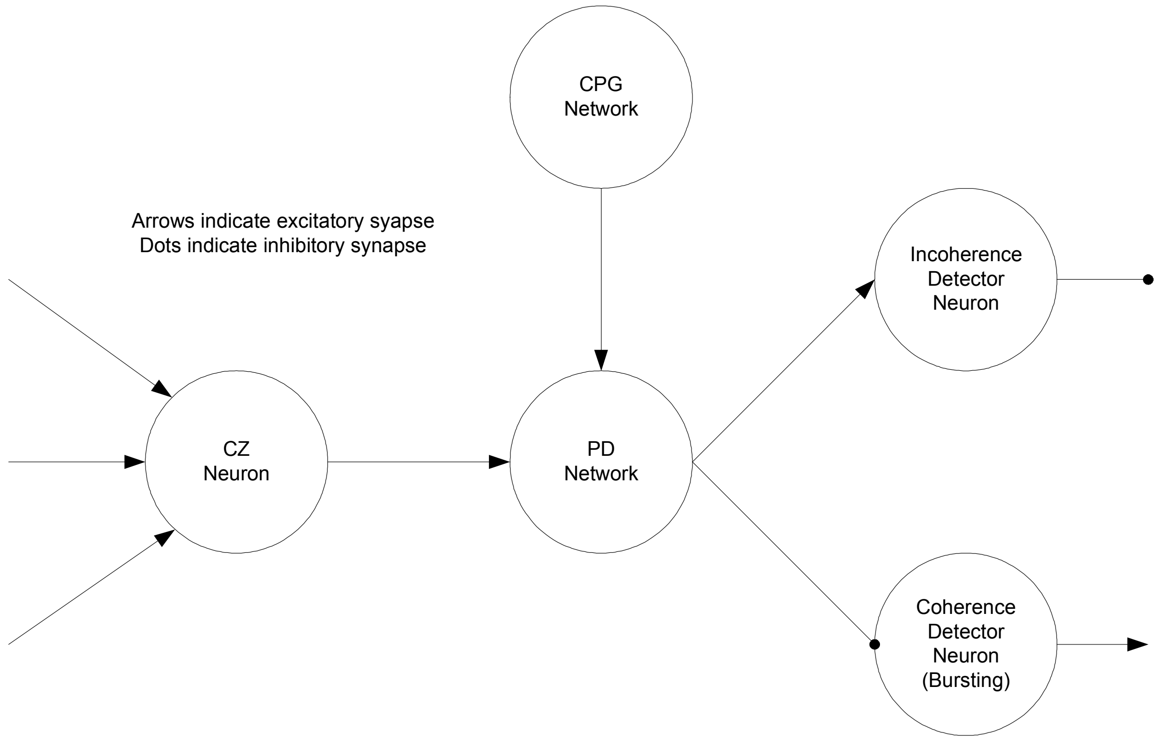


Figure 11: Schematic network for implementing coherence and incoherence detection. The central pattern generator subnet (CPG) establishes the frequency band for detection. The phase detector network (PD) fires in response to phase delay differences between the input CZ neuron and the CPG. This output acting to inhibit the firing of a bursting neuron and to excite a non-bursting neuron signals the state of comparison.

the optimization problem. Minsky and Papert proposed something like this in their idea of a "society of mind" [57]. In essence, this idea is to take one large NP-complete problem and section it into N small-scale NP-complete problems such that optimization of the system as N increases can scale as polynomial in N. The issue, of course, is "Can this be done, and if so, how?" The problem that arises when a problem space is sectioned is putting it back together again. In exchange for making problem solution easier within a section, we pick up the problem of interactions at the border line between sections. In system design this is often called the "interface problem" between blocks in a system block diagram. What goes on at the boundaries usually affects what goes on deep inside the block, and it is this facet of problem sub-space interactions that makes the sectioning problem difficult to properly solve. Merely having a "network of networks" architecture does not by itself guarantee an answer to this issue. The overall system is still *one* system and we have to deal, in a manner of speaking, with the "off-diagonal terms" in the system's "state matrix." Putting it in the language of philosophy, the problem that sectioning raises is one of dealing with what Kant called "the synthesis of unity in apperception."

What we badly need here are new ideas. One possibility is to analyze the problem in terms of required information flows within the system. A recent conference paper, not directed at neural networks but rather at information flow in distributed systems, employed an interesting analytical approach based on eigenvector analysis of the system's reachability matrix [89]. What was shown in this paper was that the eigenvector of the largest eigenvalue of a network's reachability matrix (which is merely a special case of a network's connection matrix) provides a direct node-by-node measure of information flow rates entering each node in the system (under certain implicit assumptions on entropy rates in different pathways). Furthermore, the reachability matrix for even relatively strongly recurrent network graphs tended to be sparse or at least triangular. While it is not presently clear whether or not this analysis method can be fruitfully applied to the sectioning problem for large-scale neural networks, the problem motivating this method is that of system adaptation to variable environmental conditions (particularly in robotic systems) and adaptation of variable-structure information processing systems. This general problem class is close enough to the optimization problem involved in general purpose neurocomputing to at least make the approach in [89] interesting. Furthermore, HJB optimization does not specify the "nature" of the "cost function" to be optimized. The Informax principle is arguably an adaptation scheme conforming to the Bellman principle, although not to the usual HJB equation, in which the distribution density of neuronal firing rates is maximized under a set of statistical assumptions [93]. However, despite claims that the Informax principle is "probably the simplest and most general principle" for network adaptation, Informax learning rules for PCNNs are very complicated and not at all well understood theoretically. It works by decision rules aimed at maximizing the mutual information between the input and output of the system and bears something of a resemblance to the BCM rule. Unfortunately, it is also computationally very expensive in both hardware and software forms. However, one of the principal features Informax adaptation exhibits in simulation is *disconnection* of some weights in favor of others [93]. It is here where we find a tie-point between this algorithm and the analysis presented in [89].

Under fixed connections, the response of a system will more or less quickly come to be dominated by the maximum eigenvalue of the connection matrix by a well-known theorem of Shannon's. The elements of the eigenvector of that eigenvalue tell us the magnitude of the eigenfunction response at each node in the network. In the case of both Informax and the analysis of [89], the observable is the entropy of the node's output and this can be estimated in real time from firing rate measurements [93]. Therefore there would seem to be a connection between what the Informax principle attempts to accomplish and the basic idea of [89]. However, how strong this connection may be and whether or not it is capable of producing a more efficient version of an Informax-based learning rule remains to be seen.

Purely mathematical approaches, e.g. Bellman optimization or the Informax principle, have a disadvantage that was pointed out long ago by Zadeh. When we pick some cost function to optimize, the optimization algorithm will optimize that function, but this does not automatically imply that the resulting system is "optimized" in terms of performance measures more directly representative of the purpose of the system. For example, many adaptive control systems are predicated on minimizing mean-squared error (MMSE), and optimal control strategies for doing so are known. However, it turns out that optimizing a different cost function, the integral of time-multiplied absolute error (ITAE), has been demonstrated to yield faster settling time and better disturbance rejection [94]. These performance metrics are clearly of more direct importance for what the control system is intended to accomplish and ITAE simply does better for them than does MMSE. The lesson here is plain: The cost function to be optimized should be one with either a direct or at least a demonstrated relationship to "what really matters" in assessing the "real" (intended) performance of the system.

In the case of the general purpose neurocomputer, the behavior "that really matters" in the architecture proposed here is equilibration. Equilibration is directly tied to the representation of objects, the invocation of motoregulatory expression (actions), learning behavior, and the achievement of a structured world model by the system. Cost functions that do not have a known link to equilibration, e.g. mutual information such as in the Informax principle or Hebbian dynamics in the BCM model, may well end up ill-serving this important system behavior. Equilibration implies that the degree to which the "energetics" of affective perceptions as activation regulations are minimized, and those as terminating regulations are maximized, is a more reasonable measure of "cost" and "value" in the system. *If* these can be linked to the Akuzawa-Ohnishi design indices [89], i.e. if these "costs" and "values" are operationally related to information concentration and distribution within the system, it is possible that a marriage of the hierarchical design approach presented in [89] with the Informax principle could be a foundation for adaptation when adaptation is viewed as equilibration of assimilation and accommodation. But in any event, cost and value optimization functions based directly on measures of affective perception representations have a direct tie to the *a priori* "goal" of achieving states of equilibrium in the system. Certainly their analysis is applicable for evaluating what Minsky and Papert call the "interaction and insulation" issue in network-of-network architectures [57], and "interaction and insulation" among representations is a basic character of the phenomenon of attention.

Another practical issue that will be immediately apparent to any engineer is what we might call the extensive "child development" issue implicit in this paradigm. One can easily take note of the fact that the development of cognitive and even motor control structures as called for by this model is based upon extremely general purpose elements of knowledge representations, i.e. the categories, the development of schemes based on the interaction types in the equilibration hierarchy, and so on. This model contains no rationalist "innate knowledge" of specific objects, and the innate capabilities the system does have to be provided with are on the order of very simple motoregulatory reflexes ("instincts"), fundamentally topological perception features, and a general control structure that must be designed to enable the system to develop its own schemes of locomotion and thinking based on its own experience. Such a system quite obviously has no more native "first turn on" capabilities than a newborn baby.

Equally obvious is the fact that the system's "mental development" could require a very long "education process." This is a feature that is obviously in severe conflict with the practical engineering consideration that a machine of generally widespread use must be one that can be mass produced and delivered in a much more timely fashion. The pace of technology advances and the financial undesirability of having to have any inventory- and time-intensive "school system" (populated by otherwise finished goods) for training intelligent systems are both decisively unfavorable considerations affecting the practicality of this architecture.

However, and unlike in the case of human beings, it is technically feasible for machines to be produced as "adult clones." What I mean by this is that once we have gained experience in seeing what sort of common structural and parametric features a machine develops to cope with its environment, these structural and parametric features can be *pre-incorporated* into copies of the machine. **This amounts to the hypothesis that the majority of "root" schemes, concept structures, and affective connections are likely to be of a quite universal character for any given machine in any given application environment**. What precisely these structures look like, or if indeed such a 'universal' structure really exists, is at present wholly unknown, and its discovery (or proof of non-existence) is a foundational research problem. This has several implications. First, we must design the initial machinery with *observability* features that enable us to probe and measure the variables that produce binding codes. This requirement is not unlike the technical problem of providing "test vectors" in digital ASIC designs, but the analog features inherent in the machine structure obviously make this a much bigger issue for a general purpose neurocomputer. Second, where adaptive weights are employed in the system these weights must be not only observable but also pre-settable unless non-volatile weight storage mechanisms are employed. Otherwise, the machine would revert to "new born" status every time it was powered down and restarted. Volatile weight storage mechanisms imply the requirement that the system contain one or more special purpose microcontrollers whose function is to provide: 1) an easy means for reading out the weight-setting variables; 2) an easy means of restoring these settings on power up; and 3) a practical means for communicating with the "outside world" to both effect the storage and retrieval of parameters as well as to improve the designer's ability to observe what is going on within the neurocomputer during its "cognitive development." All this must be accomplished without impacting the ability of the system to have a large number of neurons. Because the initial prototypes of a machine such as the one envisioned here is likely to require multi-chip solutions, the issue of inter-chip communication is intimately tied in to these considerations. It should be noted in this regard that an architecture based on the convergence zone scheme provides naturally for "funnel points" where a large number of signals merge down into a much smaller set in the feed-forward pathway. It therefore seems likely that system planning will be most effective if partitioning among chips takes place at the point of convergence zone feed-forward connections.

In conclusion, what this paper has attempted to provide is a general framework for posing the problems we must solve, and hence the word "preliminary" in its title. Solutions to the general issues just presented must be researched. This will involve better and more quantitative problem statements and the undertaking of specific, focused research projects aimed at providing answers to these questions. Project proposals most likely to produce the key advancements in knowledge that are needed are those which can restate in mathematical terms the qualitative problem statements given here. However, in formulating these research plans an eye must always be kept on the big picture presented here because it is this general paradigm that provides the context for contributions made by these researches.

VI. References

1. W.S. McCulloch and W.H. Pitts (1943), A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophysics*, **5**: 115-133.
2. J. von Neumann, *The Computer and the Brain*, 2nd ed., New Haven, CN: Yale University Press, 2000.
3. A.W. Burks, H.H. Goldstine & J. von Neumann (1946), Preliminary discussion of the logical design of an electronic computing instrument, in *Collected Works of John von Neumann*, A.H. Taub (ed.), vol. 5, pp. 34-79, NY: The Macmillan Co., 1963.
4. F. Rosenblatt (1958), The perceptron: a probabilistic model for information storage and organization in the brain, *Psychological Rev.* **65**: 386-408.
5. B. Widrow and M. Hoff (1960), Adaptive switching circuits, *1960 WESCON Convention Rec.*, 96-104.
6. D.E. Rumelhart, J.L. McClelland, and the PDP Research Group (1986), *Parallel Distributed Processing* (in two volumes), Cambridge, MA: The MIT Press.

7. T.P. Trappenberg, *Fundamentals of Computational Neuroscience*, Oxford, Oxford University Press, 2002.
8. W. Gerstner and W. Kistler, *Spiking Neuron Models*, Cambridge: Cambridge University Press, 2002.
9. D. Nykamp and D. Tranchina (2000), A population density approach that facilitates large-scale modeling of neural networks: Analysis and application to orientation tuning, *J. Computational Neurosci.*, **8**: 19-50.
10. A. Hamilton, A.F. Murray, D.J. Baxtor, S. Churcher, H.M. Reekie, and L. Tarassenko, Integrated pulse stream neural networks: Results, issues, and pointers, *IEEE Trans. Neural Networks*, vol. 3, no. 3, 1992, pp. 385-393.
11. R.B. Wells and A. Vongkunghae, A model reference neural network scheme for obtaining laser printer modulation templates, *Proc. 28th An. Conf. Indus. Electron. (IECON02)*, Nov. 5-8, 2002, Seville, Spain, pp. 1872-1875.
12. A. Vongkunghae and R.B. Wells, A method for obtaining robust modulation templates for laser printers, *Smart Engineering System Design*, vol. 13, pp. 907-913, NY: ASME Press, 2003.
13. T.W. Berger, M. Baudry, R.D. Brinton, J.-S. Liaw, V.Z. Marmarelis, A.Y. Park, B.J. Sheu, and A.R. Tanguay, Jr., Brain-implantable biomimetic electronics as the next era in neural prosthetics, *Proc. IEEE*, vol. 89, no. 7, 2001, pp. 993-1012.
14. L.O. Chua, M. Hasler, G.S. Moschytz, and J. Neirynck, Autonomous cellular neural networks: A unified paradigm for pattern formation and active wave propagation, *IEEE Trans. Ckts. Syst. I*, vol. 42, no. 10, 1995, pp. 559-577.
15. R. Wells (2003), The spinal sensorimotor system, *MRCI Tech Brief*, pts. I-IV, Moscow, ID: The University of Idaho MRC Institute.
16. A.S.I. Zinober (ed.), *Deterministic Control of Uncertain Systems*, London: IEE Press, 1990.
17. D. Wunsch, private conversation, Sept. 22, 2003.
18. J. Haugeland (ed.), *Mind Design II*, Cambridge, MA: The MIT Press, 1997.
19. K.M. Ford, C. Glymour, and P.J. Hayes (eds.), *Android Epistemology*, Menlo Park, CA: The MIT Press, 1995.
20. J. Piaget, *Genetic Epistemology*, NY: W.W. Norton & Co., 1971.
21. I. Kant, *Kant's gesammelte Schriften*, Akademie der Wissenschaften (eds.), in 29 vols., Berlin: Walter de Gruyter, 1910-1983.
22. C. Shannon (1948a), A mathematical theory of communication, *Bell Syst. Tech. J.*, vol. 27, pp. 379-423, July, 1948.
23. C. Shannon (1948b), A mathematical theory of communication, *Bell Syst. Tech. J.*, vol. 27, pp. 623-656, Oct., 1948.
24. A.F. Murray, D. Del Corso, and L. Tarassenko, Pulse-stream VLSI neural networks mixing analog and digital techniques, *IEEE Trans. Neural Networks*, vol. 2, no. 2, 1991, pp. 193-204.
25. L.O. Chua and T. Roska, The CNN paradigm, *IEEE Trans. Ckts. Syst. I*, vol. 40, no. 3, 1993, pp. 147-156.
26. A.F. Murray, Pulse arithmetic in VLSI neural networks, *IEEE Micro Mag.*, Dec., 1989, pp. 64-74.
27. B. Liu and J.F. Frenzel, A CMOS neuron for VLSI implementation of pulsed neural networks, *Proc. 28th Ann. Conf. Ind. Electron. IECON02*, Nov. 5-8, Sevilla, Spain, pp. 3182-3185, 2002.
28. B.C. Barnes and R.B. Wells, A versatile pulse-mode biomimic artificial neuron using a capacitor-free integrate-and-fire technique, *Proc. 29th Ann. Conf. Indus. Electron. IECON03*, Nov. 2-6, Roanoke, VA, pp. 2968-2972, 2003.
29. R.B. Wells and B. Barnes, Delay-resistor implementation of integrators in biomimic artificial neurons, *Proc. 28th Ann. Conf. Indus. Electron. IECON02*, Nov. 5-8, Sevilla, Spain, pp. 3186-3190, 2002.
30. R.B. Wells, Modulation channels in biomimic artificial neurons, *Proc. 28th Ann. Conf. Indus. Electron. IECON02*, Nov. 2-8, Sevilla, Spain, pp. 3209-3214, 2002.
31. R.B. Wells and B. Barnes, Capacitor-free leaky integrator for biomimic artificial neurons, *IEE Electron. Ltr.*, vol. 38, no. 17, 15th Aug., 2002, pp. 974-976.
32. B.C. Barnes, R.B. Wells, and J.F. Frenzel, PWM characteristics of capacitor-free integrate-and-fire neuron, *IEE Electron. Ltr.*, vol. 39, no. 16, 7th Aug., 2003, pp. 1191-1193.
33. H.S. Ranganath, G. Kuntimad, and J.L. Johnson, Pulse-coupled neural networks for image processing, *Proc. 1995 SoutheastCon*, Raleigh, NC, 1995, pp. 37-43.
34. J.L. Johnson and D. Ritter, Observation of periodic waves in a pulse-coupled neural network, *Optics Ltr.*, vol. 18, no. 15, Aug. 1, 1993, pp. 1253-1255.
35. J.L. Johnson and M.L. Padgett, PCNN models and applications, *IEEE Trans. Neural Networks*, vol. 10, no. 3, 1999, pp. 480-498.
36. J.L. Johnson, Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariance for images, *Applied Optics*, vol. 33, no. 26, 10 Sept., 1994, pp. 6239-6253.

37. Y. Ota and B.M. Wilamowski, VLSI architecture for analog bidirectional pulse-coupled neural networks, *Proc. 1997 Intl. Conf. Neural Netw. (ICNN'97)*, Houston, TX, June 9-12, 1997, pp. 964-968.
38. Y. Ota and B.M. Wilamowski, CMOS architecture of synchronous pulsed-coupled neural network and its application to image processing, *Proc. 26th An. Conf. IEEE Indus. Elec. Soc. (IECON'00)*, Nagaya, Japan, Oct. 22-28, 2000, pp. 1213-1218.
39. B. Widrow and M. Lehrs, 30 years of neural networks, *Proc. IEEE*, vol. 78, no. 9, 1990, pp. 1415-1442.
40. J.A. Anderson, *An Introduction to Neural Networks*, Cambridge, MA: The MIT Press, 1995.
41. Ch. von der Malsburg and W. Schneider, A neural cocktail-party processor, *Biol. Cybern.* **54**, 29-40 (1986).
42. R. Eckhorn, H.J. Reitboeck, M. Arndt, and P. Dicke, Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex, *Neural Comput.* **2**, 293-307 (1990).
43. W. Senn, Th. Wannier, J. Kleinle, H.-R. Lüscher, L. Müller, J. Streit, and K. Wyler, Pattern generation by two coupled time-discrete neural networks with synaptic depression, *Neural Comput*. **10**: 1251-1275 (1998).
44. R.J. Brown, *Adaptive Multiple-Output Threshold Systems and Their Storage Capacities*, Ph.D. Dissert., Stanford University, Stanford, CA, 1964.
45. A.R. Damasio, Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition, *Cognition*, **33** (1989), 25-62.
46. G.A. Carpenter and S. Grossberg (1987), A massively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision, Graphics, and Image Processing*, **37:** 54-115.
47. G.A. Carpenter and S. Grossberg (1987), ART2: Self-organization of stable category recognition codes for analog input patterns, *Applied Optics*, **26:** 151-162.
48. G.A. Carpenter and S. Grossberg (1990), ART 3: Hierarchal search using chemical transmitters in self-organizing patter recognition architectures, *Neural Networks*, **3:** 129-152.
49. A.R. Damasio, The brain binds entities and events by multiregional activation from convergence zones, *Neural Comput.*, **1**: 123-132 (1989).
50. W. James, *The Principles of Psychology*, NY: Dover Publications, 1950.
51. J. Piaget and B. Inhelder, *The Child's Conception of Space*, NY: W.W. Norton & Co., 1967.
52. J. Piaget, *The Grasp of Consciousness*, Cambridge, MA: Harvard University Press, 1976.
53. J. Piaget and R. Garcia, *Toward a Logic of Meanings*, Hillsdale, NJ: Lawrence Erlbaum Associates, 1991.
54. R.M. Harris-Warrick and R.E. Flamm, Chemical modulation of a small central pattern generator, *Trends Neurosci.* **9**: 432-437 (1986).
55. W.A. Woods, Important issues in knowledge representation, *Proc. IEEE*, vol. 74, no. 10, 1986, pp. 1322-1334.
56. J. Piaget, *On the Development of Memory and Identity*, Barre, MA: Clark University Press, 1968.
57. M.L. Minsky and S.A. Papert, *Perceptrons*, expanded ed., Cambridge, MA: The MIT Press, 1988.
58. B.K. Natarajan, *Machine Learning: A Theoretical Approach*, San Mateo, CA: Morgan Kaufmann Publishers, 1991.
59. J. Piaget, *Principles of Genetic Epistemology*, NY: Routledge & Kegan Paul, 1972.
60. J. Piaget, *Main Trends in Inter-Disciplinary Research*, NY: Harper & Row, 1970.
61. J. Piaget, J.-B. Grize, A. Szeminska, and V. Bang, *Epistemology and Psychology of Functions*, Dordrecht, Holland: D. Reidel Publishing Co., 1970.
62. J. Piaget, *The Development of Thought: Equilibration of Cognitive Structures*, NY: Viking Press, 1977.
63. J. Piaget, *The Construction of Reality in the Child*, NY: Basic Books, 1954.
64. J. Piaget, B. Inhelder, and A. Szeminska, *The Child's Conception of Geometry*, NY: W.W. Norton, 1960.
65. J. Piaget, *The Child's Conception of Number*, London: Routledge & Kegan Paul, Ltd., 1952.
66. L. Aiello, C. Cecchi, and D. Sartini, Representation and metaknowledge, *Proc. IEEE*, vol. 74, no. 10, 1986, pp. 1304-1321.
67. B.L. Webber, Logic and natural language, *Computer*, vol. 16, no. 10, 1983, pp. 43-46.
68. N. Kasabov, Decision support systems and expert systems, in *The Handbook of Brain Theory and Neural Networks*, 2nd ed., M.A. Arbib (ed.), Cambridge, MA: The MIT Press, 2003, pp. 316-320.
69. B. Inhelder, H. Sinclair, and M. Bovet, *Learning and the Development of Cognition*, Cambridge, MA: Harvard University Press, 1974.
70. J. Piaget, *The Origins of Intelligence in Children*, Madison, CN: International Universities Press, 1952.
71. R. Wells, *The Critical Philosophy and the Phenomenon of Mind*, in preparation.
72. J. Piaget, *Intelligence and Affectivity: Their Relationship During Child Development*, Palo Alto, CA: Annual Reviews, Inc., 1981.
73. R.S. Lazarus, On the primacy of cognition, *American Psychologist*, vol. 39, no. 2, 1984, pp. 124-129.
74. R.S. Lazarus, Cognition and motivation in emotion, *American Psychologist*, vol. 46, no. 4, 1991, pp. 352-367.

75. R.B. Zajonc, Feeling and thinking: Preferences need no inferences, *American Psychologist*, vol. 35, no. 2, 1980, pp. 151-175.
76. S.I. Greenspan, *The Growth of the Mind*, Reading, MA: Addison-Wesley, 1997.
77. S.I. Greenspan, *Intelligence and Adaptation*, NY: International Universities Press, 1979.
78. A.R. Damasio, *The Feeling of What Happens*, NY: Harcourt Brace & Co., 1999.
79. A.R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain*, NY: Avon Books, 1994.
80. S.R. Palmquist, *Kant's System of Perspectives*, Lanham, MD: University Press of America, 1993.
81. R.A. Brooks (1991), Intelligence without representation, in *Mind Design II*, J. Haugeland (ed.), Cambridge, MA: The MIT Press, 1997.
82. J. Piaget and B. Inhelder, *The Psychology of the Child*, NY: Basic Books, 1969.
83. J. Piaget, *Judgment and Reasoning in the Child*, Totowa, NJ: Littlefield, Adams & Co. 1966.
84. C. Peterson and B. Söderberg, Neural optimization, in *Handbook of Brain Theory and Neural Networks*, 2nd ed., M.A. Arbib (ed.), Cambridge, MA: The MIT Press, 2003, pp. 822-827.
85. F.L. Lewis, *Optimal Control*, NY: John-Wiley, 1986.
86. J.D. Baum, *Elements of Point Set Topology*, NY: Dover Publications, 1964.
87. A.G. Barto, Reinforcement learning, in *Handbook of Brain Theory and Neural Networks*, 2nd ed., M.A. Arbib (ed.), Cambridge, MA: The MIT Press, 2003, pp. 963-968.
88. A.G. Barto, Reinforcement learning in motor control, in *Handbook of Brain Theory and Neural Networks*, 2nd ed., M.A. Arbib (ed.), Cambridge, MA: The MIT Press, 2003, pp. 968-972.
89. K. Akuzawa and K. Ohnishi, Design indices for information connection in decentralized system, *Proc. 29th An. Conf. Indus. Electron.*, Roanoke, VA, Nov. 2-6, 2003, pp. 2417-2422.
90. D. Hsu, M. Figueroa, and C. Diorio, Competitive learning with floating-gate circuits, *IEEE Trans. Neural Netw.*, vol. 13, no. 3, 2002, pp. 732-744.
91. F.J. Pelayo, E. Ros, X. Arreguit, and A. Prieto, VLSI implementation of a neural model using spikes, *Analog Int. Circ. and Signal Process.* **13** (1/2): 111-121, 1997.
92. J.L. Johnson, Pulse-coupled neural networks, *Adaptive Comput.: Math. Electron., Opt., SPIE Crit. Rev.*, vol. CR55, pp. 47-76, 1994.
93. J. Feng, Y. Sun, H. Buxton, and G. Wei, Training integrate-and-fire neurons with the Informax principle II, *IEEE Trans. Neural Netw.*, vol. 14, no. 2, 2003, pp. 326-336.
94. D. Graham and R.C. Lathrop, The synthesis of "optimum" transient response: Criteria and standard forms, *Trans. AIEE, pt. II: Applications and industry*, vol. 72, Nov., 1953, pp. 273-288.