# The Cognitive Self-Structuring Connectionist Machine

BY

Ahmed A. Fadol

A Thesis Presented to the

DEANSHIP OF GRADUATE STUDIES

**KING FAHD UNIVERSITY OF PETROLEUM & MINERALS**

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

**MASTER OF SCIENCE**

**In**

**COMPUTER ENGINEERING**

May 2001

UMI Number: 1406097

UMI®



Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by Ahmed A. Fadol under the direction of his thesis advisor and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE IN COMPUTER ENGINEERING.

Thesis Committee

Thesis Advisor

Member

Member

22 May 01
Department Chairman

Dean of Graduate Studies

23/5/2001
Date

عمادة الدراسات العليا
DEANSHIP OF GRADUATE STUDIES

# ACKNOWLEDGEMENT

Acknowledgement is due to the Department of Computer Engineering and Center for Communications and Computer Research at King Fahd University of Petroleum & Minerals for supporting this research.

I wish to express my appreciation to Professor M.K. Ibrahim who served as my major advisor. I also wish to thank the other members of my thesis committee Professor M. Abd-El-Barr and Dr. A. Bouridane.

TABLE OF CONTENTS

LIST OF FIGURES

# THESIS ABSTRACT

Ahmed A. Fadol
The Cognitive Self-Structuring Connectionist Machine
Computer Engineering
May 2001

A cognitive architecture for artificial intelligence is proposed. The architecture is cognitive in the sense that fundamental cognitive hypotheses served as guidelines for its design. It is termed self-structuring because of its ability to autonomously extract and provide structure to what would otherwise be considered unorganized data. In addition, it is considered a connectionist machine in view of its establishment on biological principles and its distributed storage of data. The architecture is unique in that it incorporates many of the desirable features of both symbolic and connectionist systems, while being entirely based on proven biological and psychological knowledge of the human brain, thereby providing a sound foundation for future development. The architecture is exemplified by the construction of an artificial vision system based on the basic machine, and the scope for further enhancement and development is outlined.

ملخص الرسالة

الاسم: أحمد عباس فضل

عنوان الرسالة: المكنة الإدراكية البانية ذاتياً الاتصالية

التخصص: هندسة الحاسب الآلي

تاريخ التخرج: صفر ١٤٢٢ هـ

تقترح هذه الدراسة معمارية ادراكية للذكاء الاصطناعي. المعمارية تعتبر ادراكية لإن الفرضيات الإدراكية والاحساسية استخدمت كوسائل إرشادية في التصميم. تسمى بانية ذاتية بسبب قدرتها على استخراج وتكوين البنية ذاتياً بالنسبة لبيانات تعتبر غير منظمة. بالاضافه الى انها تعتبر مكنة ارتباطية في ظل تأسيسها على مبادئ بيولوجية، وتخزينها التوزيعي للبيانات. وتعتبر هذه المعمارية فريدة من حيث إحتواءها على الكثير من المميزات المرغوبة لكلاً من النظم الرمزية والارتباطية، مع كونها مؤسسة كلياً على المعارف المثبتة البيولوجية والنفسية عن المخ البشري، وبذلك توفر اساس ثابت للتطوير المستقبلي. المعمارية ممثلة ببناء نظام بصري اصطناعي مبني على أساس المكنة المبدئية، ويطرح مجال التحسين والتطوير المستقبلي.

# CHAPTER ONE

# INTRODUCTION

Artificial Intelligence (AI) as a formal discipline has been studied for nearly sixty years [1]. It has been marked by long periods of incremental work within established guidelines, and the branching of novel sub-fields that occasionally generate original areas of research. An important observation is that all sustained bodies of work depend on similar philosophical and technological assumptions of the period that greatly influence their development and many times inadvertently lead to their constraint.

Early as 1941, the idea of using computers to play chess was discussed and theoretically developed using the min-max search algorithm [1]. During the 1940's and early 50's the discipline of cybernetics, although with a fundamentally different approach, had many of the goals currently defined as those of artificial intelligence. The late 1950's and early 60's witnessed great efforts in the development of distributed linear threshold computational devices commonly known as perceptrons [1], [2]. The 1960's saw the solidification of AI research into two main approaches: symbolic and connectionist methods, which was a factor in the slow down of research into alternate approaches [3]. The late 60's and early 70's witnessed the great influence of von Neumann computers and the Lisp programming language in the proliferation of serial search-based artificial intelligence research [2]. These two influences together with some early successes and the publishing of the 1969 book "Perceptrons" that carefully analyzed the limitations of the one-layer perceptron, led to the sharp decrease in connectionist based research for the next fifteen years [2]. The mid 80's witnessed a resurgence in connectionist research after the development of the back propagation algorithm and the availability of suitably powerful computing devices [1].

Some advocates of the bottom-up "learning by example" paradigm mainly used by connectionists have proposed that the best way to achieve intelligent behavior is to start with the simulation of insects, and gradually progress with the simulation of more complex creatures, until ultimately human intelligence can be simulated. The answer to this proposal is the Chimpanzee Counterexample given by Sowa [4]. The chimp brain is the closest to the human brain with similar structure and about one-third the neurons. Chimps are comparatively as perceptive and agile as humans, and remarkably, three-year old chimps can outperform three-year old humans in non-verbal IQ tests. However, that is where the similarities end. Chimps cannot develop even rudimentary levels of symbolic reasoning due to the inexistence of the basic enabling structure. In contrast, enabling structures are highly advanced in humans, and are required for such tasks as symbolic reasoning and theorem proving, among the other problems that require true intelligent behavior. Therefore, it can be seen that intelligence simulated for insects or even chimpanzees, cannot be considered sufficient or even a first-step for the task of constructing a system with the simulation of human intelligence as its ultimate goal.

What is required in constructing a system that is desired to ultimately simulate human intelligence is a comprehensive look at the evidence discovered thus far describing human intelligence. Since the available evidence is quite sparse and often conflicting, any attempt at designing such a system must make diligent use of the available evidence and apply it in the most suitable manner. Many formerly constructed systems did not make efficient use of the available information about human intelligence during their design. While the reasons for this are varied, usually they were caused by the constraint of available simulating resources, the particular inclination of the researchers, or loosing sight of the ultimate goal and optimizing the systems for more immediately useful applications at the expense of further studying and refining the architectures.

In designing this proposed system, it was decided to base its design on relevant evidence pertaining to human intelligence, beginning at the fundamental neural operations and interconnections of the brain to better understand how it is constructed and finishing with the results of experimental psychology to better understand what are the end results of the brains operation. With this information at hand, the ultimate task was to explain and simulate how the brain works. While the available work on this topic is still insufficient for fully understanding the workings of human intelligence, by accumulating the results of a broad range of research approaches all attempting to ultimately explain intelligence albeit at different levels, basic guidelines and mechanisms of intelligence emerge. Through the utilization of these emergent insights into human intelligence, an abstract system was designed and constructed.

To demonstrate the efficacy of this system on common tasks requiring intelligence, it was decided to apply it to vision. Human vision has been the subject of considerable study by the intelligence research community and it is reasonably well understood compared to other areas of intelligence. Therefore, besides helping to demonstrate the operation of the system, the mechanics of vision will also prove to be useful in understanding human intelligence as a whole. In designing this system, the ultimate goal was to explain human intelligence and the design was not garnered to any particular application or constrained by the available means of simulation. It must therefore be stressed that although the system is being applied initially to a visual application, it is also suitable for the simulation of other areas of intelligence, providing that they are well understood and synchronized to requirements of the system.

# CHAPTER TWO

# INTELLIGENCE AND VISION

## 2.1 DEFINITION OF INTELLIGENCE

In order to discuss intelligence, the term itself must be carefully defined. Many times intelligence and cognition are implicitly assumed to cover a much wider scope of powers, which are considered to be in the dominion of mind. Listed below are the dictionary definitions of these terms that will facilitate their correct identification and subsequent discussion [5].

| | |
|---|---|
| Intelligence | The capacity to acquire and apply knowledge.<br>The faculty of thought and reason.<br>Superior powers of mind. |
| Cognition | The mental process or faculty by which knowledge is acquired.<br>Something that comes to be known as through perception, reasoning or intuition; knowledge. |
| Mind | The human consciousness that originates in the brain and is manifested especially in thought, perception, feeling, will, or imagination. |

> The totality of conscious and unconscious processes of the brain and nervous system that directs the mental and physical behavior of a sentient organism.

From these definitions, the impression is given that intelligence and cognition are only components of mind, and that they can be simulated and explained independently from other aspects of mind. Cognition complements intelligence in that it explains the processes by which knowledge is acquired and processed, both terms operate on knowledge states, and since knowledge states are particular instances of information states, the underlying process is considered to be information processing [3]. Henceforth, the term intelligence will be used to encompass cognition and the desired meaning will be implicitly understood from the context.

The separation of intelligence from other components of mind is quite controversial and is usually met with two opposing perspectives. The first perspective is that different components of mind are simply emergent properties of complex agents with knowledge states, and that once these agents reach a certain level of complexity, components such as emotion and subjectivity will simply appear. On the other hand, the second perspective is that knowledge state processes cannot exist without considering all other components of mind. This perspective would indicate that only an agent, which from the onset encompasses all components such as subjectivity, emotion, etc. can be considered intelligent [3]. Distinguishing intelligence from mind is crucial in the development of artificial intelligence systems. Trying to simulate the entire mind from the very onset is impractical due to the sheer scope of such an undertaking. Therefore, while artificially intelligent systems may not initially posses other components of mind, they are still important stepping-stones upon which future development can take place. Once a system is properly designed, regardless of its initial abilities, it will facilitate the addition of other components of mind.

## 2.2 NATURAL INTELLIGENCE

Natural intelligence depicts all forms of biological intelligence, from the most sophisticated level found in humans, down to the most primitive forms found in simple creatures. While the ultimate goal is to understand human intelligence, its overwhelming complexity and ethical constraints on experimentation place immense obstacles in obtaining even rudimentary understanding. On the other hand, while it is easier to experiment on less sophisticated creatures, their facility for explaining behavior is quite limited [6], which could result in the incorrect generalization of results to humans. Therefore, until new experimental techniques are found, or the behavior of animals can be readily explained, our understanding of intelligence must come from a careful combination of both human and non-human sources of evidence to ultimately achieve the original goal of understanding human intelligence.

In the following exposition of natural intelligence, existing evidence at the biological, physiological, and psychological levels will be discussed and analyzed in order to form a suitable groundwork for which current approaches to artificial intelligence can be evaluated and proposed architectures can be based upon.

## 2.2.1 BIOLOGICAL EVIDENCE

The human brain is dominated by the cerebral cortex, whose size and complexity in humans compared to lower mammals is presumed to be the primary source of perceptive superiority [7]. It is composed of several regions, each with a distinct responsibility such as speech, visual information analysis, or motor activity and all with largely similar internal anatomical organization. Together these regions work in concert to compose intelligent behavior, through what is thought to be a general-purpose style of computation [8], [9], [10]. This abstract organization should provide the basic framework for any biologically inspired AI system.

The elemental building blocks of the nervous system are nerve cells or neurons, which are interconnected through synapses. While being extremely diverse in terms of their shapes, sizes and the tasks for which they are optimized, they are usually composed of the same three main components: *dendrites* that consist of connections from other neurons, *cell body* that supports the functions of the neuron, and *axon* that carries the output signal to other neurons [7]. Dendrites are responsible for receiving signals from other neuron's axons via synapses. Their physiological specifications such as diameter and length are responsible for determining how the incoming signals are processed. The cell body supports the entire structure by providing basic life processes, and may even perform some information processing itself. The axon conducts signals to other neurons via synapses, and although there is only a single axon per neuron, it has the ability to branch tremendously to other nerve cells. Signals consist of chemical transmitters, which when crossing the synaptic gap affect the operation of the target neuron [7].

Synapses form the interconnections between neurons, they combine electrical phenomena and the actions of complex micromolecules in determining the dynamic activity of neurons [7]. A synapse accepts the arriving action potential from the cell body through an influx of calcium ($Ca^{++}$) ions, which causes the modulated release of a transmitter substance to the receiving cell that binds to its receptor molecules and causes a change in its electrical activity, which influences the electrical signaling to the receiving neuron [9]. Despite the large and varied number of synapses in the brain, it is generally accepted that they can be broadly classified into two types according to the effect of their transmitter: excitatory and inhibitory [8]. In actual operation, the receptor molecules of the synapse accept only certain transmitter molecules. If subsequent binding occurs the configuration of the receptor molecule changes and it will ultimately define the actions of the transmitter upon the dendrites of the destination cell.

The principal task of a single neuron is the summation and temporal integration of the incoming excitatory and inhibitory signals from the synapses [7]. While, the activation potential of a neuron is nearly continuous, the interaction between neurons with action potentials is clearly not [8]. In addition, the size and shape of the dendrite tree and its branches has critical influence on how the signals are processed. The ultimate result of processing may lead to the activation of the neuron and a resultant signal propagates through the axon to other neurons.

A powerful characteristic of the nervous system is its ability to grow and develop [7]. Dendritic and axonal branches grow in length and their branches increase through development. In fact, humans are born with only 1.5% of the synaptic connections found in the typical adult [11]. In addition, when also considering that synaptic densities in humans are far greater than those found in any other biological brain, there seems to be a direct relation between the level of intelligence and the number and complexity of the connective branches and synaptic connections.

In general, it can be said that the biological model of intelligence relies on a single abstract model of computation, utilizing a generic model of building blocks and operations. However, in order to optimize for specific applications, the building blocks are modified to better suit the tasks required. Therefore, any biologically correct artificial model of intelligence should have a single basic mode of computation, a single configurable building block framework and the ability to accommodate operations of varying complexity among its inner components.

### 2.2.2 PHYSIOLOGICAL EVIDENCE

A single entity can be represented over a pattern of activity distributed over many elements. Each element is then involved in representing many entities in an efficient way, which utilizes the processing abilities of a network of simple neuron like elements [2], [10]. A great motivation for such a representation is that its strengths and weaknesses closely match those of the human brain. It is important to keep in mind that such a representation is not in conflict with the extensive evidence of localization of function within the brain. It is not even an alternative to abstract representation techniques used heavily in traditional artificial intelligence. However, distributed representation provides an implementation technique of these larger functional components, while allowing for the emergence of powerful properties such as content addressable memory, generalization and the ability to deal with imperfect data.

The ability of humans to deal with incomplete and sometimes even partially faulty data when accessing their memories, is an example of content addressable memory that is extremely beneficial but difficult to implement on conventional computers. Distributed representations could allow for such content-based memory access [2]. This would be done through the partial activation of some units that would represent "microfeatures" whose activation would subsequently result in the activation of other units and result in the fetching of stored data. The significance of such a technique is that an entity is not localized in any particular location and is only recovered through the existence of correctly activated connection paths between the relevant units. While this method should be able to recover stored data upon correct input, it may well happen that certain inputs will recover the wrong stored data or even a combination of several stored items of data. This characteristic is also found in humans where erroneous inputs may lead to the recovery of "incorrect" data or many times a fusion of many different items of data. In view of this evidence, any artificial system should be able to intrinsically generalize and infer relationships among its newly introduced and previously stored data.

In storing a new item of data previously stored items must not be erased or critically distorted, this must be achieved through the meticulous adjustment of existing connection paths between units upon the introduction of any new item. A possible solution is to store all items through orthogonal patterns of activity. While this would entirely prevent any interference between items, it would also eliminate the property of generalization, which is a useful phenomenon that helps in dealing with situations that are similar although not identical to previously experienced situations. Since humans are very good at generalizing newly acquired data with previous knowledge, the importance of generalization as a component in any intelligent system is further demonstrated [2].

Neurons in the brain are extremely slow when compared to conventional computing components functioning at speeds in the milliseconds or even 10's of milliseconds. This fact when combined with the reality that highly sophisticated perceptual processes can be performed in a few hundred milliseconds, means that the number of serial steps is limited to only about a hundred or so, and is clear evidence of massive amounts of parallelism occurring in the brain [2]. However, the brain cannot be classified simply as a parallel processor. Due to the processing of temporally spaced signals it is more correctly classified as a spatiotemporal processor [7].

Neurons themselves are functionally very simple when compared to conventional computing components. They exist in large numbers, in the order of $10^{10}$ to $10^{11}$ in the brain. Each neuron is an active component that can have from 1,000 to 100,000 synapses on their dendrites and a similar number as inputs to dendrites of other neurons. Neurons are generally not activated by a small number of inputs. This suggests that computation is a statistical process in which the decisions are the product of a large number of units that offer continuous information. Reliability is also a direct consequence of such statistical behavior in that the degradation of a single neuron will not result in the collapse of the system's performance, rather a graceful degradation will occur as more and more neural units are destroyed. Statistical operation and reliability of intelligence can be attributed to the internal operational mechanisms and the immense number of neurons, thereby requiring any simulation of intelligence to similarly contain a similar statistical mode of operation working over a large number of individually insignificant operational units. Another possibility for neuron activation is that its input dendrites interact nonlinearly [8]. This may be probable if voltage dependent channels in the dendrites that would non-linearly transform the signals are proven to exist. The decentralized nature of the brain's structure also means that there is no central executive overseeing operation. This means that all parts of the brain work cooperatively and contribute to the overall performance, this feature is well characterized by higher level functions of the brain. Communication among neurons is done only by excitation and inhibition through the connections, which means that only very small amounts of actual data are transferred relating information on whether to excite or inhibit the connection.

Since memory storage and information processing are essentially performed by the same "circuitry" in the cerebral cortex, this means that they must therefore be closely related and use essentially the same type of computation [8]. This observation is quite significant in view of the many artificial systems that separate data storage and processing into distinct modes. Physiological conformity requires that data storage and processing be treated similarly and use a homogeneous mode of computation that suffices both their requirements, alleviating the distinction between storage and processing altogether.

The connection topology of neurons in the brain is interesting in that most of the connections are very short, while the minority are long. This means that there are spatially distinct regions in the brain, where nearby regions of the brain map into other nearby regions. It is also important to point out that the different regions of the brain are interconnected in significantly different ways [8]. The reason for this localization may be to conserve connection length, because of the very large number of connections resulting in a high-dimensional space that must be embedded into a relatively small three-dimensional space. This results in units with the highest degree of interaction located closest to each other [2]. Connections among the operational units must therefore be dynamically configurable and expandable in order to allow for the maximum amount of flexibility and efficiency during operation and not be statically predefined.

### 2.2.3 PSYCHOLOGICAL EVIDENCE

The essential core of the constructivist view of cognition is that the observation of our world must involve some type of conceptual framework [12]. Since all that is learned from our surroundings must reach us through our senses or through language, therefore from the onset any stimuli is tainted with the particular conceptual bias of the method of reception. Any system that attempts to represent the type of conceptual structures used by humans must be able to represent two distinct types of hierarchy: *part to whole*, where the whole is equal to the sum of its parts and *type to instances of type*, which specifies that known properties of type must be inherited by its instances [2].

The idea that memory is physiologically distributed within constrained regions in the brain is now a generally accepted notion [8]. Even so, the question of how to represent memory within the brain; either as specific or general information has caused a significant debate. A large number of experiments have revealed that humans seem to extract what is common among a set of experiences thereby storing this information as abstract concepts while neglecting insignificant features, this would suggest that human memory stores general information. However, an increasing number of experiments are also proving the significance of specific stimuli in many tasks that involve abstraction, thereby emphasizing the importance of specific information.

An approach, which attempts to describe this model is the *enumeration of specific experiences view*, it simply says that each object or stimulus is in itself an extracted concept that is used abstractly. This approach while solving the memory representation dilemma by discarding with the notion of general information, requires unlimited memory space and a correspondingly sophisticated search mechanism to deal with such data. While some empirical results have proven encouraging, a major criticism with this model is that it completely does away with any abstraction of data and the construction of theoretical hierarchies that are known to be in use by humans. At the other extreme, models that make use of only abstract models and no specific stored data, have not performed well. This is attributed to clear evidence showing the significant effect of particular stimuli in the behavior of abstract models, thereby opposing any premise of this model. A more probable response to this representational duality is that abstract rule-based or concept-based behavior can be attributed to processes that make use of specific events or rules, and that both general and specific information must be seized and correlated together in the same model. Therefore, the ability to deal both explicitly and abstractly with data is an important ability of the human brain and consequently an important requirement in any simulation of intelligence.

## 2.3 VISION

Since "seeing is intelligence" [13], the following discussion on vision will be central in amassing the different components of natural intelligence as applied to an actual biological system, while providing additional insight into the naturally utilized mechanisms and assist in the resolution of several conflicting operational observations.

In the following analysis of the requirements of vision, a top-down approach will be pursued, in reverse of the naturally occurring process. Starting first with vision at the psychological level, in order to better understand what are the crucial requirements of vision and what are supplementary. Then, while gradually removing any supplementary details of vision and discussing further implementation specific details that are required for this primitive vision system, a final representation will be reached at the neural cell level.

## 2.3.1 LEARNING AND UNDERSTANDING

A single object can have an infinite number of possible projections upon the retina. Orientation can vary, ultimately giving rise to different two-dimensional projections. The object may be partially occluded, missing some components, under moderate levels of visual noise, or even be a simplified line drawing. With rare exception, an object can always be readily classified by humans either as an instance of a familiar category, or as an unfamiliar instance of a new category, which itself is a form of classification [14].

Less-complex animals such as insects rely almost completely on an unlearned, innate observation of objects, and their ability to show perceptual learning is almost nonexistent. Such as the case with bees, they do not have to learn about flowers, they simply search for nectar where their ancestors found nectar, it can be said that the pattern of flower petals is some how built into the bee's brain. At the other extreme, a human child is born with a colossal capacity for learning, but an extremely limited knowledge of its surroundings. Except for some innate abilities such as perception of dangerous depths, as shown by the "visual cliff" experiments [15], and the ability to direct its eyes to objects of interest, which may be to boot-strap the learning process, little else is innately known. These two examples show that all life forms regardless of their sophistication have some intrinsic knowledge, and that the capacity to learn and develop from sensory perception is what distinguishes highly complex creatures from less complex life forms. Consequently, any artificial human visual system, must have a firmly incorporated mechanism to help in acquiring new knowledge, thereby providing the ability to learn new objects.

In describing any scene consisting of various objects, any proposed description technique must allow for the ability to match, compare and contrast between different scenes [16]. Even more important than simply analyzing a scene is the ability to learn and draw conclusions from its analysis, which in turn, can lead to the creation of more complex abstractions. Essentially, since a human will always serve as the scene analysis system's mentor, it is reasonable that the same relational lexicon be used, such as left-of, right-of, and top-of. While simplifying interaction, this will also help in using what is known psychologically about human intelligence in the development of the system. Even though these relations are quite abstract, they may still be too discrete. For example, people have great difficulty in keeping track of the faces of a six-colored cube, if it is required that they roll it around in their mind and are required to describe the relations [16]. The use of more abstract lexicon such as next-to and opposite-to would however simplify the relation task when dealing with spatial rotation. Experimentally, it has been discovered that two main types of connections exist for the visual pathway [6]. The first is horizontal providing for lateral interaction among cells, and the second is vertical carrying information from the photoreceptors to the brain. This is an indication to the relative simplicity of associative mechanisms implemented in the human brain and reference to a possible relationship between association at the abstract level and neural level. These two types of connections may also give a biological basis to the rudimentary type to instance of type and part to whole conceptual structures used by humans.

Therefore, in any simulation the description of association should be kept as simple as possible and avoid any excessive complexity.

It is now widely believed that complex objects are internally represented symbolically as a collection of related simpler objects, through both psychological experiments and machine simulations [8], [14], [17], [18], [19]. A proposed general theory, Recognition-by-components (RBC) [14], is based on the fundamental assumption that a modest set of components ($N \leq 36$) can be derived to readily contrast the properties of edges in a two-dimensional image, which can represent a pseudo three-dimensional image. This theory is analogical to speech recognition, where only about 55 phonemes (primitive speech elements) are required to represent virtually all the words to any spoken language [14].

The RBC theory assumes that when an object is viewed it is segmented into separate components at regions of deep concavity, particularly cusps. The resulting segmented objects are then approximated by the volumetric components, and related to each other to recognize an object. It is worth mentioning that this theory describes primal access only, and does not take into account secondary descriptors such as color and texture. The proposed RBC primitive components differ in curvature, collinearity, symmetry, parallelism, and cotermination. These components are referred to as non-accidental, because they are unlikely to be a consequence of an accident of viewpoint, which results in optical illusions.

In a further increase in granularity of the RBC approach, the use of codons as simple primitives to describe plane curves was proposed [14]. These descriptors would further dissect primitive object shapes, into curves that are represented using basic primitive codons. This method may be considered as an implementation of the RBC technique that simplifies the task of representing complete shapes.

While the two preceding methods effectively explain how complex objects can be broken down into primitives, such as volumetric components or codons. They do not explain why some objects are recognizable almost immediately, while others require more thought. For example, when learning the alphabet of a new language, each letter is painstakingly scrutinized with regard to the edges, cusps, dots etc. in order to recognize it. At this stage, the letters are completely dissected into the most primitive components for recognition. As proficiency is gained, combinations of components are immediately recognized (the similarity between R, P, and B for example). Finally, the complete letter is stored as a "primitive" object. Ultimately, different combinations of these letters are also stored as "primitive" shapes. Therefore, it seems as though when trying to recognize an object we first check the primitive stored shapes, before attempting to delve deeper into the more primitive components such as edges and cusps.

### 2.3.2 RECOGNITION

The time available for the human brain to perform complete visual processing of an image (estimated at 200-300 milliseconds), places severe restrictions on the types of algorithms that could be used in the cerebral cortex. This time restriction excludes the possibility of any complex cooperative algorithms that require the extensive exchange of strict numerical information among neurons through iteration [14]. However, it has been suggested that through the use of probabilistic coding that would represent the probability of firing, enough time may be available for a cooperative algorithm to converge and reach a state of "relaxation". At the other extreme of relaxation algorithms are "one-shot" algorithms that are able to converge in a single pass. In viewing these two algorithms as extremes of a continuum of strategies used by the cortex, any strategy may be stylized depending on the situation [8]. For example, one-shot algorithms may be used for the recognition of objects through experience, and relaxation algorithms may be used in the recognition of more complex or novel objects.

Extracting objects from an image generally consists of isolating the foreground from the background, by pinpointing changes in contrast. In processing the line of contrast of any object, the amount of attention placed just outside the line of contrast (separating it from the background), is as high as that placed just inside the line of contrast [20]. This dual treatment of neighboring contrast regions doubles the ability of the visual system in segmenting objects from their backgrounds. In addition, there may also be memory-based circuitry that tries to bias the visual processor's contrast tolerance to interpret parts of the field as objects and the rest as background. Through object contours, we are able to detect the edges and cusps that are vitally important in object recognition. Contour detection is a continuous process, the highest contrast contour is detected first, followed by lower levels of contrast that further help in detecting the object's details, such as texture, or curvature.

As mentioned previously, color and texture are not essential for initial recognition of objects, resulting in only the highest contrast contour being essential. In addition, not all components of an object need to be present for primal access [14]. Therefore, the task of recognizing any complex object is ultimately a matter of identifying the arrangement of a limited number of critical components. It is known that human memory is biased toward the regularization of irregular shapes, according to RBC, this is because perceptual input is mapped onto a representational system based on a limited set of regular primitives. It was also proven experimentally that complex objects that are composed of a number of simpler shapes could be recognized with only a partial complement of components (Fig. 2.1).

Objects can often be more easily recognized from some orientations compared to others. In some cases, an object may not be recognized at all if viewed from an unfamiliar orientation. The RBC explanation is that unfamiliar orientations depend on the recognition of components normally not associated with that object, consequently the detected components do not readily match the components normally related to the object, and recognition becomes more difficult or even impossible. It is also important to point out that many times an object being viewed from a different orientation, resulting in the segmentation into a set of components that greatly differs from its exemplar view can be recognized as the same object. For example, when viewing a car from the side and front, almost no common components exist, but the object is still recognizable if the views have been previously learned and associated with the class of "car". Therefore, any system must be able to associate sets of components that may be completely orthogonal to the same object class.

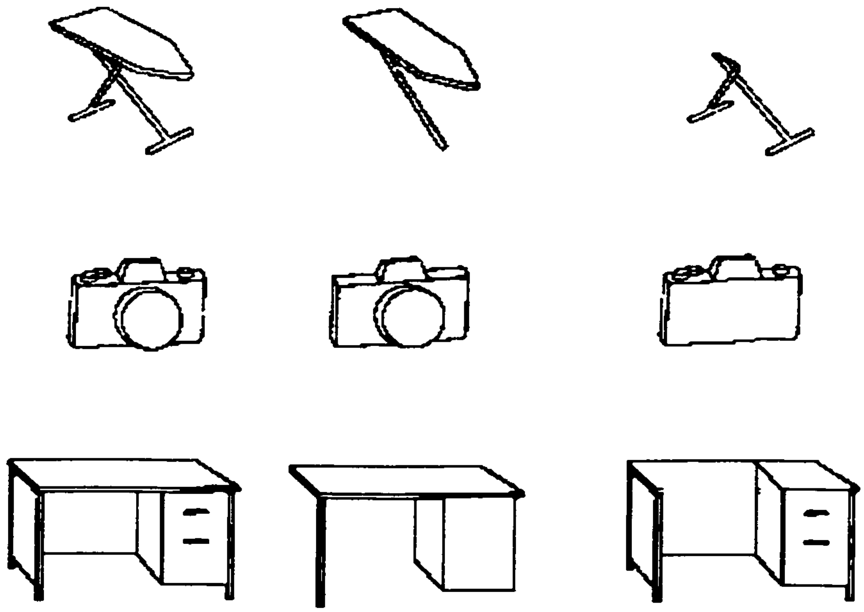

Figure 2.1 Recognition of a complex object with only a limited number of components
Complete objects are shown in the left column, as can be seen in the middle column recognition is still possible with the removal of non-critical components, while recognition becomes much more difficult with the removal of critical components in the right column [16].

An object is parsed into components at regions of concavity [14]. The degree of resistance to noise in the recognition of an object, depends on the particular areas that have been affected and to a lesser extent the total amount of noise. If noise degrades the edges of an object, it is still visually recognizable even with relatively high levels of degradation. However, if the noise affects the cusps of the object, the degree of recoverability is seriously damaged (Fig. 2.2). This phenomenon, highlighting the greater importance of cusps in regard to edges has also been observed experimentally [14]. Its explanation as given by the RBC hypothesis, is that the distorted edges can be "filled-in" through the non-accidental relations of collinearity and curvilinearity, and any cusps that are missing would lead to the loss of an entire component, since they define the end-points for edges.

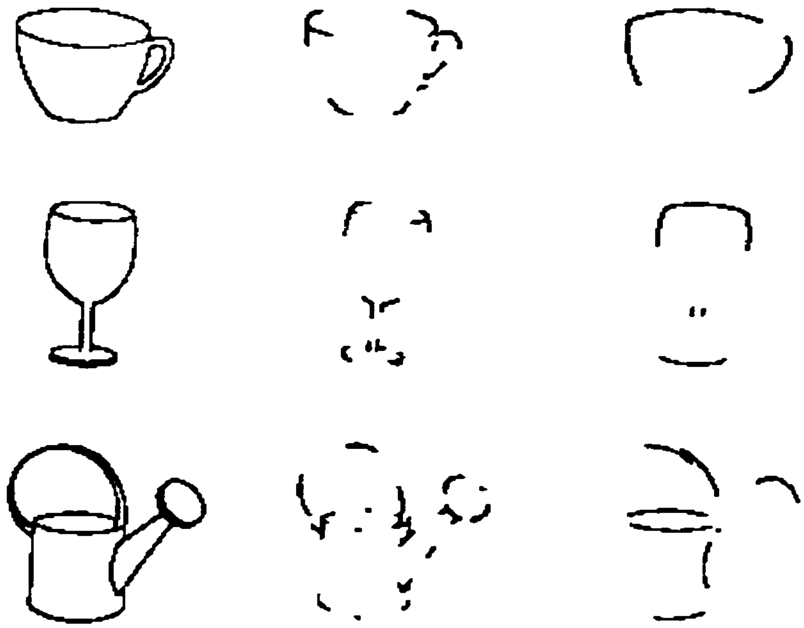

Figure 2.2 Importance of cusps versus edges
The left column shows the complete image, while the middle column shows the objects with noise affecting the edgesand the right column shows the objects with noise affecting the cusps [16].

In a series of experiments first performed by G.M Stratton [11], [15] to test rotational invariance of the visual system, humans and animals were fitted with inverting glasses. Using the knowledge that the retinal image of the outside world is represented upside down (from basic properties of lenses discovered by Keppler), a series of experiments were conducted with the goal of discovering whether adults (both human and animal) could learn to internally see the world right side up after a life time of seeing it upside down. While humans were able to adapt to the new situation within a few days (some even able to ski within a week), animals showed almost no signs of adaptation even after several months. The results of such experiments are highly debatable. On one hand, it seems as though humans retain visual plasticity into adulthood. Alternatively, the results could be interpreted as though learning consisted of a series of quick specific adaptations overlying the original perception, rather than a complete reorganization of the original perceptual system. The results can even be attributed completely to positional plasticity of the human body, which is only calibrated by vision [11]. Since all the experimental subjects were immediately able to adjust to their natural viewing orientations, this rules out the hypothesis of complete perceptual reorganization. As to the remaining two hypotheses, they are both plausible but extremely difficult to prove or disprove. The relevant outcome from these experiments is that biological vision cannot be considered rotationally invariant, since not even humans were able to immediately adjust to the effects of the inverting glasses.

In a series of discoveries first started by Hubel and Wiesel [20] relating to the limits of rotational resistance, it was found that many neurons in the visual cortex of cats or monkeys are sensitive to the direction of a line of contrast on the retina. Through experimentation, it was found that neurons have a maximal response to at least a dozen preferred directions. This means that objects that lie on these preferred directional fields will get maximum response from the visual system, and may suggest the existence of directional components in the visual cortex that enable an efficient means of extracting information and generating a suitable contour description. These directional components may give evidence to the existence of an expensive rotation algorithm, which cannot meet the recognition time requirements except through division of the task. There is also reason to believe that the visual processor provides only a limited amount of rotation and aspect invariance [20]. Of course, we can immediately recognize an object in all rotational positions but only after having seen it in a variety of rotations. It seems as if there is an approximate $\pm 40°$ limitation to rotational-invariance, after which it becomes considerably more difficult to immediately recognize the object [20]. In any case, rotation invariance is a very expensive operation. This is probably why the retina makes use of multiple directional components to distribute the massive amount of computation, and even so, is quite limited in function. Therefore, a simulation of the visual system needs only to be resistant to rotation and not require complete invariance. On the other hand, there seems to be no such limitation for scale-invariance. When an object is changed only in scale, the cusps remain constant in terms of their directional angles and only the edges change in length. This phenomenon can be considered to further corroborate the importance of cusps over edges in object recognition.

## 2.3.3 REPRESENTATION

The question of how biological visual systems represent objects internally for recognition is a debated issue in the vision research community. Among the hypotheses suggested is the *reconstructionist* paradigm [16], which holds that objects are represented internally to a certain degree as three-dimensional analogs of the actual physical objects. The opposing hypothesis suggests that representations are viewer-centered and largely two-dimensional, this hypothesis has reemerged following both computational and experimental support after it was initially rejected on philosophical grounds. Techniques have been proposed to represent a three-dimensional object model from a linear combination of several two-dimensional views, utilizing only a small set of corresponding features or a linear mapping of the entire view [21], [13]. In addition, there is general agreement that vision is symbolic to a certain extent, and a naive observer will most likely relate objects symbolically (next-to, opposite-to, etc.) rather than through a discrete description [16]. Thus, the symbolic nature of relations discards the notion of dimension in imagery, since all images are essentially dimensionless fragments that are related to one another and the difficulty in providing discrete descriptions of objects is likely due to the lack of an underlying three-dimensional representation. Therefore, any visual system only requires a set of interrelated two-dimensional views to adequately describe a physical object.

In converting the incoming image into frequency-coded electrical signals, the retina makes use of electro-optical non-uniformly distributed transducing neural receptors

[6]. Less complex life forms depend most greatly on motion detection nerves, and their limited set of visual nerves is quite complex in terms of diversity. However, as the complexity of the life form increases, the diversity of the nerves decreases in correspondence to an increase in the total number of nerves [20]. Life form complexity peaks in humans, where the majority of receptors consist of cones and rods, with an order of magnitude more rods than cones. Cones are responsible for day vision, thereby providing a high degree of acuity and the ability to perform color processing. Night vision is provided by slow-responding rods, which are light sensitive and achromatic. This distribution would explain the ability of humans in making unhindered use of achromatic images and the crucial importance of contrast in visual processing over all other visual details.

Almost all information transmitted by neurons is coded into action potentials, whose presence or absence relays information to other neurons [8]. Early experiments were able to correlate the action potential of single cortical neurons to simple features of sensory stimuli [8]. This gave special importance to the cellular level of information coding, and gave clear evidence that single neurons were able to code simple sensory features and perhaps other simple percepts. However, analysis of an entire image, which is a combination of simple features, requires the concerted effort of the entire visual system. It is unlikely that a highly complex object is stored in its entirety inside a single neuron, even if it is a frequently used "primitive" object. It is well known that each visual element of a complex object is represented by a node that corresponds to a particular set of feature values, thereby leading naturally to the necessity of a distributed representation

[14]. The more likely method of dealing with primitive objects is that the components of the object (the cusps and edges) become solidly connected thereby giving the illusion that it is a single object, hence speeding recognition.

## 2.4 CONCLUDING REMARKS

The first objective of this chapter was to identify and analyze the different components of natural intelligence at the biological, physiological and psychological levels. Through this identification, the essential requirements of any intelligence system are described. The second objective was the analysis of visual processing at its different stages, this helps in the characterization of how the components of intelligence are actually implemented in a real application. These minimal requirements of human intelligence and vision will be used to evaluate existing AI architectures in the following chapter and form the basis of the proposed AI architecture.

# CHAPTER THREE

# ARTIFICIAL INTELLIGENCE

## 3.1 APPROACHES AND GOALS

The field of artificial intelligence can simply be defined as the attempt to understand human intelligence, the construction of systems that perform tasks when performed by humans are considered intelligent, and informally to pass the Turing test [1].

The answer to the following question largely defines the general approach taken by AI researchers: Can intelligence be characterized abstractly as a functional capability that is merely realized by biological organisms? If so, the study of the biological brain and the constraints of human psychology are unnecessary for simulating intelligence. This thoroughly functional view of intelligence is held by many *symbolists*, where there is no claim that knowledge inside an agent is internally represented in explicit form or that its processes have any inferential basis. However, *connectionists* consider intelligence deeply associated with underlying biological systems. Many of their abstract architectural proposals are based on the information processing and smooth concept learning believed to happen in the brain [3]. In contrasting these two approaches, it can be said that they both actually extract some concepts from biological phenomena in designing their abstract architectures, albeit at different levels. With connectionists choosing a more profound view of biological processes to base their work upon, and symbolists choosing to make use of only the functional results of biological systems such as how humans use knowledge to reason and achieve goals. This incomplete consideration of the complete process of natural intelligence by both approaches is reflected directly by their complementary strengths and weaknesses.

It may seem at first that symbolic AI techniques are fundamentally different from connectionist techniques and that they can never be truly combined into a single system. One has to consider the fact that they are usually implemented on the same type of computer and described using the same set of high-level languages. Therefore, from the theoretical foundations of computer science they are not irreconcilably different, regardless of whether they use complex serial operations on lists of symbols, or primitive parallel operations on numbers [3].

Both AI approaches belong to the same scientific foundation of natural intelligence and they are simulated using the same computational theory. Therefore, it is not theoretically impossible to design and implement a system that encompasses the complete range of natural intelligence and at least match, if not exceed their combined strengths.

## 3.2 ANALYSIS OF THOUGHT

The relationships between conscious thoughts in humans have provided a great stimulant to current practices in AI [3]. It is known that in general, thoughts sequentially follow one another. While thinking for a particular purpose, thoughts are directed, through the process of acceptance, rejection, and continuous focusing until the purpose or goal of thinking has been reached. This process can be referred to as *deliberation*, in humans it can last for several seconds or indeed much longer when trying to reach an ultimate solution such as in the case of problem solving. In contrast to deliberation, short-term thought processes, which last in the order of milliseconds are referred to as *sub-deliberations*, natural-language understanding and visual perception are the most common examples of this type of thought [3].

The reasoning view of deliberation considers architectures that are closely governed by the rules of logic and are designed to simulate the logical relations thought to be in used rational thought [3]. Attempts have been made to achieve such architectures. The first method is based on the use of logic machines that work on large sets of knowledge represented in logical formalism and use logical rules as their primitive operators. The alternative second method uses machines that generate thoughts that may not be necessarily logically correct, after which correct logical patterns are applied to verify the conclusions reached by the machines thoughts. However, an important design issue is the topic of control. When a logical rule should be applied and how subsequent inferences are derived are all topics of control of which logic itself provides no guidelines. This is why control is usually task-specific and must be explicitly modified for each application. Consequently, such architectures generally are difficult to build and do not posses high flexibility. An alternative view of deliberation does provide a rudimentary control mechanism as part of its architecture and is termed the goal-subgoal deliberation framework [3]. The premise of this technique is the belief that goal thoughts spawn sub-goal thoughts recursively until all the sub-goals are solved and ultimately the final goal is solved. This deliberation technique resembles to a large extent searching in a problem space. Proposed solutions to sub-goals are verified through logical rules that are not operators of the architecture itself, but rather used as pieces of knowledge. Therefore, these logical rules can be considered context and domain independent.

Unlike deliberative architectures, which focus on the rational solution of complex problems in its effort to understand intelligence, proposed sub-deliberative architectures focus on the study of seemingly irrational solutions to problems [3]. For example, upon visual recognition of a person after a long period of time, it is difficult to pinpoint the actual visual parameters that triggered recognition. It is believed that such sub-deliberative architectures are a direct reflection of neural components of the brain, and therefore have a biological basis, which may be the key in understanding intelligence.

The most significant direct offshoots of deliberative and sub-deliberative architectures are symbolism and connectionism respectively, with the main division between these two techniques being the manner in which they consider and process information. However, since humans are capable of both types of thought, they must both be included in any faithful simulation of human intelligence. In addition, with many common perceptual scenarios it is difficult to find a clear-cut distinction as to the type of thought being used. Therefore, it can be said that thought is not polarized into two techniques, but rather that they are the extremes of a single band. This artificial polarization of thought may only be the result of the underlying limited outlook into the theoretical foundations of AI from which they are implemented. The consideration of the entire natural intelligence process may alleviate this polarization, by considering thought as a continuous band in parallel with the process of natural intelligence.

## 3.3 SYMBOLIC

In characterizing the difference between people and machines, it is frequently said that people are "smarter" than machines. Even though people are not as quick or as precise as machines, they are substantially better at perceptual problems such as scene analysis or natural language processing. The reason for this disparity in performance as explained by classic artificial intelligence (symbolists) is that there is missing "software"; that if we had the correct computer program, the unique characteristics of human information processing could be captured [2].

Symbolic artificial intelligence (SAI) was originally inspired by the problem-solving search systems of early artificial intelligence research [1]. It places significant importance on the concept of appropriate representation and its premise is that once a problem has been adequately represented, the problem is almost solved [22]. In attempting to solve a problem SAI relies on the use of a set of rules directly derived from the notions of deliberative thought and reason in finding a path between an initial state and a goal [1]. In reality, what happens is the movement from one set of expressions (symbols) to another set, where there is no known polynomial algorithm that will directly find this path [3]. This functional view of information processing is not based on any underlying biological principles of the brain. A valuable feature of such architectures is the fact that their basic operational units employ verbally recognizable concepts in their operation [3]. This information representation greatly facilitates the human analysis and understanding of such systems, and is a result of the belief that mental processes might essentially be similar to information processing as in von Neumann type machines. A particular concern is that these representations are too slow and complex for the nervous system to actually use and the dissimilar relational lexicon is extremely intricate and rigid. In addition, it is completely unlike the memory distribution model and homogeneous storing/processing employed in the human brain, it is therefore not inherently capable of generalization and robust behavior. In fact with traditional localized semantic network representations extra processes must be invoked to allow activation to spread from a local unit to neighboring units that represent similar concepts in order to perform generalization, therefore this feature is not a characteristic of the underlying architecture.

Among the most popular and basic representation methods are semantic nets, which convey meaning through nodes that denote objects, links denoting relations between objects, and link labels denoting particular relations. Feature-based object identification of an object is then done by describing it through a suitable representation, then attempting to find a suitable stored semantic net description that provides a satisfactory match. This representation method can also be extended to describe object relations and transformations, which is useful in the analysis of more abstract concepts such as language processing. The ability of symbolic systems to represent hierarchical logical concepts functionally as employed by humans is quite advanced and results in the ability to perform complex deliberation.

One of the most significant advances in symbolic AI was the development of the *schema* in the mid 1970's, to help represent complex relations that exist implicitly in the human knowledge base [8]. Therefore, schemata can be thought of as generalized concepts used to model underlying structures of the outside world. A particular dilemma, is the fact that schemata are designed to be highly structured in order to capture the regularities of situations and support the resultant inferences, while it is desired that they be highly flexible to adapt to new situations and scenarios of events. A generalization of semantic nets is the frame representation method. Apart from having a different graphical representation, the most significant enhancement to basic semantic nets in this method is the addition of the concept of inheritance, where knowledge is shared and is more easily distributable and updateable throughout the system. Among the recent modifications to frame based systems, are systems which exploit the concept of inheritance in utilizing case-specific knowledge [23], allowing the fine-tuning of domain knowledge to better match specific cases.

Learning in traditional symbolic AI is generally viewed as searching through a defined hypothesis space that grows exponentially with the size of the problem to be solved [24]. The predominant learning paradigm is referred to as "explicit rule formation" or top-down learning [1], [25]. The fundamental idea is to formulate explicit rules (heuristics) relating to a domain theory that capture powerful ideas in a concise manner, thereby avoiding exhaustive and impractical searches in large problem domains. The generation of these rules usually requires a complete and correct account of the application domain and in many cases it also requires a starting set of prepositional representations [2], this is usually extremely difficult and sometimes impossible [26]. The requirement of narrow explicit domain theories results in the system's performance not being able to degrade gracefully as the boundaries of the application domain are approached, which is in conflict with the robust operation of the biological brain [28].

Upon learning a new concept the establishment of new units or the modification of current connections may be required. Essentially, this means that such systems are able to learn in as quickly as one pass, which is primarily due to the sophisticated logic formation abilities of the system. However, the solution to the learning problem is a discrete decision that will determine whether to modify the existing connections or find a new unit with suitable connections for the learned concept [2]. A major problem with such a learning paradigm is that available units may not be used effectively and result in the inefficient use of available resources, this problem can also be attributed to the localized data representation scheme.

The traditional AI technique has been criticized with over-idealizing content and the distortion of the actual form in which knowledge really emerges [3]. Connecting inputs and outputs representing the real world have also been largely ungrounded, and implicitly assumed to exist, whereas in the human brain the entire system is seamlessly contiguous [1]. In addition, the task of specifying initial knowledge states and ultimately designing systems that are capable of learning from primitive or non-existent knowledge states poses extremely complex problems both in terms of specification and time, that are only hampered by the negligence of micro-structure. The source of the problem is that symbolic systems are locked into a fixed representational base of primitives [4], therefore a system architecture is usually not easily configured to deal with a different perceptual task from which it was originally designed.

## 3.4 CONNECTIONIST

In describing the reason for human superiority in dealing with perception-based problems as compared to machines, symbolists claimed it was a matter of having the correct "software". However, the alternative to this hypothesis is that it is a matter of having the correct "hardware". Connectionists claim that the basis for human superiority is the existence of a parallel-distributed computational architecture that is capable of simultaneously considering many pieces of imperfect or ambiguous information and constraints in the brain that is more suitable for dealing with perspective type problems than the simple machines we now have [2].

Connectionist artificial intelligence (CAI) methods draw their inspiration from what is known about real neurons, their biological make-up, how they are connected, and how they transmit information. Such systems termed as sub-symbolic are quite varied in interpreting the biological brain model and are in general parallel-processing systems that involve interactions between connected units and rely on relatively simple cooperative computations resulting in excitatory or inhibitory signals to other units for the majority of their information processing [3]. Of the most significant approaches in trying to achieve such systems are artificial neural networks (ANNs). This class of architectures was originally designed to simulate a small subset of the most prominent characteristics of the brain, but has now departed dramatically from its original biological inspirations, most notably exemplified by the back-error propagation paradigm [7].

The mode of operation is to output a correct response to every input the system receives. This input-output behavior is a function of the network architecture, where every function is computed by the individual nodes using parameters such as the connection weight [3]. This configuration is commonly referred to as a *constraint network* in which each node represents some hypothesis, and each constraint represents a constraint between hypotheses [8]. Therefore, if hypothesis X exists whenever hypothesis Y exists, there would be a strong positive constraint between them. On the other hand, if X and Y cannot exist simultaneously, they would be connected through a strong negative constraint. Likewise, all intermediate states of existence between X and Y would have corresponding values of constraint. The operation of such a network in performing information processing would be to satisfy as many of its constraints as possible through the activation of units, while giving priority to the strongest. Finally, the network would settle into a state of *relaxation* that maximizes its *goodness* value through the use of various hill-climbing heuristics. This mode of operation is incapable of one-shot learning due to its simplistic processing abilities and is quite limited in its usefulness.

Artificial neural networks follow a regular layered topology from the onset and the interconnections between these layers are usually fully connected, this contrasts with the dynamic development of interconnections in the biological model [7]. In addition, while processing of incoming signals into the neuron is relatively simple, it is not quite as simple as the summation and thresholding used by ANN's. On the contrary, in biological neurons relatively complex biochemical processes occur involving a detailed microstructure thereby enabling much more complex behavior [7].

The basic units of information storage are usually not recognizable as comparable equivalents to familiar concepts, rather they store only small details or microfeatures, which on their own are unintelligible, but whose presence or absence in groups might lead to the construction of familiar verbal concepts [3]. With this biologically faithful distributed representation, generalization automatically occurs upon introduction of new data through the modification of connection strengths of all similar activation patterns.

A unique aspect of the way knowledge is stored in connectionist systems, is that it is not stored remotely and fetched upon processing, rather it is an integral part of processing itself similar to the homogeneous storage/processing model used in the human brain. Since knowledge is effectively the strengths of the connections between the units, learning simply becomes a case of finding the correct connection strengths so that the correct activation is produced from the right circumstances through tuning its interconnections to correctly identify the interdependencies between activations [2].

Unlike top-down learning, the goal of learning in these models is not the production of explicit rules, rather to allow the model to reach a convergent state through its simple hill-climbing technique. In addition, instead of using complex learning mechanisms, only simple strength modulation mechanisms are utilized. In general, bottom-up learning can be divided into two distinct paradigms: *associative learning* and *regularity discovery* [2]. In many instances the functional difference between these two paradigms is not clear-cut, however the goals of each paradigm are usually quite distinct.

In the case of associative learning, the system learns to produce a particular pattern of activation in response to another pattern of activation. If the pattern is associated to itself, it is referred to as an auto-association paradigm and the goal is usually to self complete itself upon being presented with an incomplete version of itself. However, if distinct patterns are associated together relations between the units of the associated patterns are constructed with the goal of storing the patterns so they can be re-evoked in the future, this is facilitated through the use of an external teacher (supervised learning). Regularity discovery systems learn to respond to particular patterns of activation through classification into different categories and are only concerned with the meaning of a single response to the entire input. No external teacher is required in such systems, as they have internal teaching mechanisms built-in, commonly referred to as unsupervised learning. However, the categories can be explicitly defined or the system is to independently discover them through its use of internal feature representations. Despite the differences between the two paradigms they essentially perform one task and that is pattern matching. This simple operation alone is incapable of producing logical operations that even approach those found in humans, although there has been some limited success in reproducing the behavior found in lesser creatures, which are incapable of complex symbolic thought.

In learning any novel concept, all that needs to be done is modify the connections between the present units to create a new pattern of activity for the new concept. The advantage of this is that no new units and connections need to be established upon learning, but the problem of finding a suitable activation pattern that will require the least amount of modification to the weights so that it will not excessively disrupt existing patterns and at the same time introduce positive effects to the existing network [2]. Therefore, the solution of this problem is central to the development of such learning techniques. A significant problem with this paradigm is its complete ignorance of any learning theory [28]. As such, performance is usually corrupted by the generation of spurious connections that may involve irrelevant features. In addition to the inability to construct complex features from the initial simple features, which may lead to more sophisticated learning techniques or even simplify the task of learning itself are not performed [26].

Many connectionists argue that the mind's biological mechanisms cannot be simulated with artificially pre-labeled constructions as in symbolism and that just because certain pieces of knowledge are used in deliberation they are not necessarily analogous to what we have in our consciousness [3]. Even though human cognition has a sequential feel in the process of deliberation and going from state to state, connectionists argue that it is still not implemented sequentially. If an attempt is made to actually model these states and the microsteps inside them, the simplest task of cognition would require an enormous number of microsteps if performed sequentially. The addition of any constraints would only increases the time required by a sequential machine, whereas in humans processing becomes quicker as more constraints are presented [2]. However, it is important to point out that the majority of connectionists do not rule out the existence of a sequential macrostructure of cognition in the same way that the study of subatomic particles does not deny the existence of interactions between atoms, in other words connectionism describes the internal structure of larger units, which compose the cognitive process.

Artificial neural networks excel at many of the tasks for which symbolic AI are deficient at: pattern recognition, learning and generalization [4]. For the task of pattern matching, the patterns themselves are not stored, rather the connection strengths between the units are stored, which allows the patterns to be recreated. While this allows for highly effective sub-deliberative operations, it is a theoretical challenge for connectionists to demonstrate that other more complex problems such as logic problems can be based on the same architectures used for much simpler tasks [3].

The lack of structure in connectionist systems is the dominant criticism, since it clearly exists in human intelligence and is a requirement for any plausibly intelligent system [2], [4]. An essential component of the ability to adapt to change, is the ability to isolate and manipulate knowledge and information, without any type of conceptual framework to express data, this is not possible in typical neural network models [4]. Among the proposed solutions is the utilization of multi-level networks that would perform pattern recognition at multiple levels and use the resulting patterns of each level to partially constrain the results of the next level, this could simulate logical thought by steering solutions through multiple levels until an ultimate solution is reached [3]. However, even with the use of multiple levels, the most basic forms of structure that are critically important in the simulation of human intelligence are still non-existent.

Another criticism to connectionist systems that is also applicable to all biological brains is that their operation for the most part is completely unintelligible to humans. The progress of such systems is usually untraceable, the information structures stored in the weights of its hidden layers are inaccessible, and no underlying abstract computational level theory is available to corroborate any results [4]. The lack of human-understandable knowledge representations in present connectionist systems causes concern about the credibility of the solutions reached [24]. Even though the alleviation of this criticism would facilitate the development of connectionist systems, its existence would not interfere with any broad requirements of intelligence.

## 3.5 HYBRID SYSTEMS

The complementary nature of symbolic and connectionist architectures has been exploited in the construction of hybrid architectures that attempt to embody the desirable features of both architectures by using information from one source to offset missing information from another source. Hybridization is usually attempted by two general approaches [27]. The first approach is *integration by transformation*, where the knowledge base and the reasoning methods are transformed into a representation suitable for the other paradigm. The alternative approach is *integration by cooperation*, where both paradigms retain their original functionality and are used cooperatively to solve a particular problem.

Among the most significant recently proposed integration by transformation architectures is KBANN (Knowledge-Based Artificial Neural Network). This architecture attempts to insert and translate hand-constructed symbolic rules into an ANN that is then refined using standard learning algorithms and results in a refined domain theory that is later extracted [28]. The major difficulties faced thus far with this architecture are its limited rule syntax, insufficiently tight symbolic and neural learning, and unsatisfactory rule extraction mechanisms [26].

The main source of improvement to this architecture over a standard ANN is the ability to train faster because of better starting points specified by the inserted symbolic rules. The architecture is still not inherently able to support hierarchical structure, does not provide a satisfactory mechanism to support multi-typing of weights, and is a further departure from the connectionist biological origins. Therefore, this architecture can only be said to be an improvement to existing ANN structures, as compared to a genuinely hybrid AI architecture.

## 3.6 CONCLUDING REMARKS

The goal of this chapter was to discuss the present realizations of artificial intelligence and their adherence to the previously discovered requirements of natural intelligence. In viewing the two current paths of AI research in terms of their original inspirations, functionalities and abilities, it becomes quite clear that their selected design approaches resulted in their subsequent deficiencies. By not basing their designs on the entire spectrum of natural intelligence, important design-dependent capabilities are lost.

The result of this work will be to demonstrate that the complete consideration of evidence into human intelligence from various viewpoints will help in its ultimate identification and explanation. This will lead to the alleviation of the current AI dichotomy, by demonstrating a system that is innately capable of performing both types of applications normally associated with symbolism and connectionism, in addition to new applications that require a blend of their capabilities. By not simply adapting an existing paradigm to acquire some of the desirable properties of the opposing paradigm, a truly complete system free of any inherent design problems will be constructed. The proposed system will also demonstrate how these different viewpoints, while initially seeming to be dissimilar, in fact, corroborate each other at different levels.

# CHAPTER FOUR

# COGNITIVE SELF-STRUCTURING CONNECTIONIST MACHINE DESCRIPTION

## 4.1 DESIGN OBJECTIVES

The main objectives in designing this system were to incorporate both the unique complementary features of symbolic and connectionist AI systems and more importantly the intermediary perceptual abilities that require measured combinations of their abilities into a single homogenous system through the careful consideration of the entire process of natural intelligence.

Many previous attempts at the construction of hybrid AI systems made use of well-established connectionist or symbolic systems adapted to encompass additional features from the opposing class of systems, such as the use of symbolic data in training neural networks [28] or the ability of generalization added to symbolic systems [8]. While these hybrid systems do have additional functionality, they still do not possess all the desired features of both types of AI systems, inadvertently acquire some of the undesired characteristics of the opposing class, and move still farther away from what is known about biological intelligent systems. In addition, even if a hybrid system was constructed that was able to combine the abilities of both symbolic and connectionist architectures, it would still not suffice the requirements of human intelligence. For example, perceptual abilities such as empirical and mentor learning are only extremes of a continuous scope of learning used by humans [29]. Therefore, current AI architectures are ultimately limited, either alone or as hybrids. The solution to this problem may be to base the design of any system that attempts to faithfully simulate human intelligence on the same underlying structural framework used by humans in order to encompass the complete scopes of basic perceptual processes.

It was decided that this proposed system should be cleanly designed from beginning, avoiding the connectionist/symbolic AI dichotomy altogether and more importantly to have the design adhere as closely as possible to the established discoveries of the human intelligence system and its underlying structure. While the verified discoveries are still lacking, deficiencies in certain areas can be made up by discoveries in other areas. It was decided that the system should be primarily based on the biological and physiological discoveries of the brains neural system at the low-level, and the proven psychological experiments that hint at the high-level functionality of the brain. The largest gaps in understanding the human intelligence system were encountered for describing its mid-level functionality. This functionality pertains to "how" the low-level neural structure is able to perform the high-level perceptual tasks. Therefore, the main challenge was to try bridging these gaps with what is well understood or has been deemed highly probable in regards to the information processing performed at the neural level resulting in the perceptual functionality.

## 4.2 DESCRIPTION

The Cognitive Self-Structuring Connectionist machine (CSSC machine) is a system that consists of two elemental mechanisms: cells and links. Through the complementary nature of these two mechanisms, the system is constructed. What follows is an abstract description of the system's main components (Fig. 4.1). These components are the main information processors of any application based on this system, although they must be complemented by application specific components to efficiently perform the task required. This is reminiscent of the human brain where different regions are optimized for different tasks with special application specific neural cells [7].

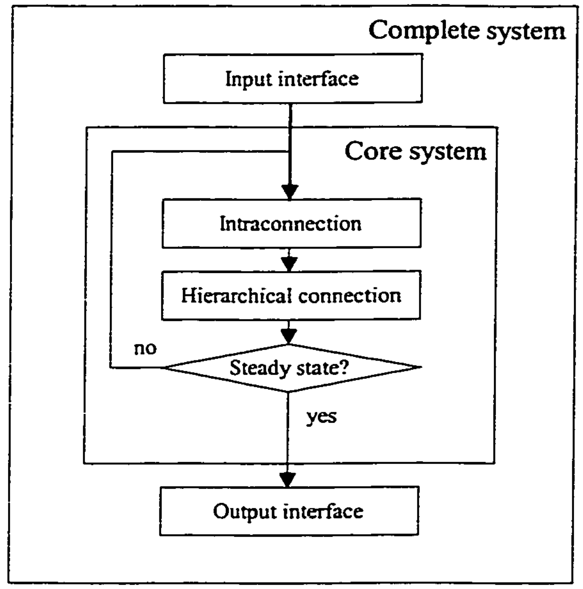


Figure 4.1 Overview of system configuration and operation

## 4.2.1 CONNECTION CELLS

Connection cells (c-cells) form identical autonomous units that cannot act independently of their environments. Therefore, they will only act in response to an external form of excitation provided by connection links. An appropriate action is taken based on the state of the incidental links and the result of their operation could lead to the generation of new links that will then act upon other cells. These cells are designed to act as data processors rather than data holders; therefore, they only retain the minimum amount of information (retained from incidental connection links), and purge or propagate any additional data from the incidental links to the created links.

Every c-cell cycles through three different states of involvement: *dormant*, *inactive*, and *active*. Initially every cell is dormant, it will change to the state of active when it is accessed by an incidental link and perform its required processing based on the nature of the links. Once processing is complete, the c-cell changes to a medial inactive state, which means that it is part of the current activation path, but its current processing task has been completed. However, the c-cell is still a candidate for re-activation if it is later found suitable. When a c-cell has been activated, it will be in one of two structural generation modes: *part to whole* or *type to instances of type*, which will be referred to as intraconnection and hierarchical operations respectively. When a c-cell is in a particular structural mode, it will regulate the types of links that can connect to it and the types of links that can be generated by it (Fig. 4.2).

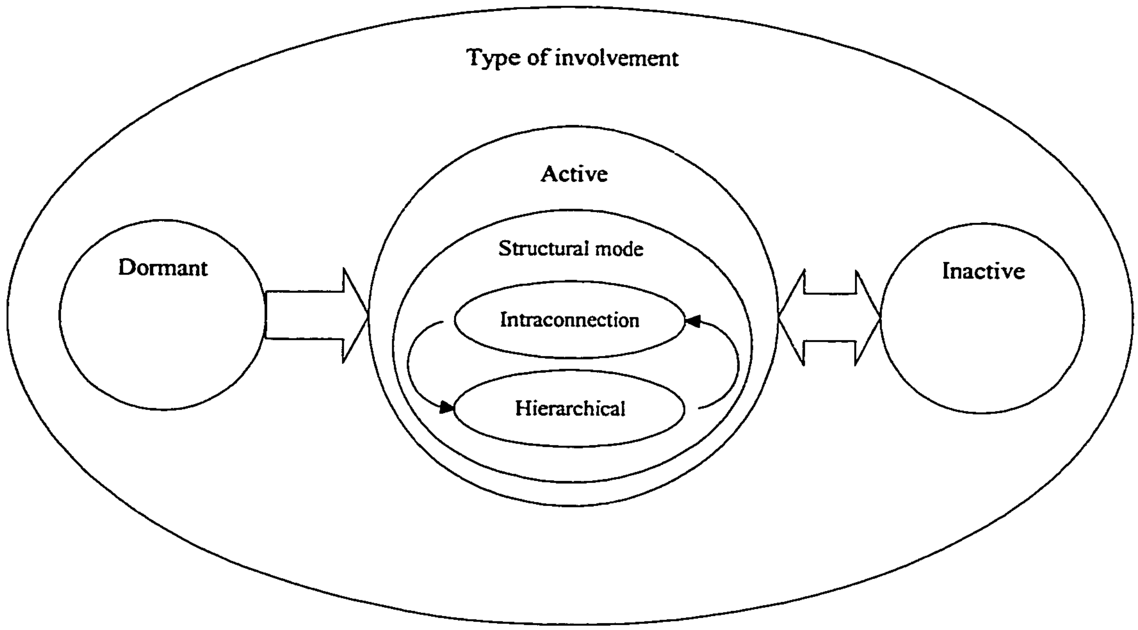


Figure 4.2 C-cell states of involvement

A c-cell oscillates between these states of involvement and can only enter structural mode when it is active.

When a link is incidental to a cell, it will insert an entry into the cell in order to establish association with the cell, transfer required information and facilitate matching of future links to correct cells.

While active, the action of the c-cell depends on its structural mode, and any data held within it; the state, type and number of incidental links, and their corresponding data values. The structural mode of the c-cell will dictate what will cause it to activate/ generate links, and what the destinations of these links will be. If the c-cell is in intraconnection mode, it will attempt to connect to other c-cells that are in the same involvement state and structural mode. The particular connection topology will depend on the data held within the involved c-cells that has been placed by previous links. After all intraconnection mode c-cells have completed their operations, their mode becomes hierarchical. That is, after the operation of part to whole has been completed, it is time for the type to instance of type operation. In this mode, all the "parts of the whole" will attempt to link to a type c-cell. This type c-cell is most probably dormant in its involvement, and is not in any structural mode. Upon activation, its structural mode becomes intraconnection. As can be seen, c-cells continually oscillate between the two structural modes until system-wide processing reaches a steady state.

In biological correspondence, c-cells abstractly represent the simple cell body with the embedded link entries representing synapses. Relatively complex processing operations that result in link generation are similar to those of biological synaptic operations. This offers a much closer association with biological systems than most other connectionist-based systems especially artificial neural networks. In addition, an underlying structural theory for link generation exists. This theory is inspired by the basic conceptual structures used by humans at the psychological level and correspondent with the necessity of an organized physiology of neural interconnections. The introduction of structure into neural interconnections is a major improvement over other connectionist-based systems and will be explained fully when discussing connection links.

A cell will for the most part deal with imperfect or incomplete data coming from its incidental links, therefore activation, is based upon attaining certain thresholds, which can be raised or lowered depending on the degree of rigidity required in the decision making process. Multiple thresholds can be established depending on the requirements of the application. Therefore, the potential complexity of operations is much higher than in simpler connectionist systems, while not approaching the complex data representations in symbolic systems. In most cases, they will relate to two types of link-based information: the total number of incidental links that are active or dormant and the aggregated weights held by the active links. Since a single link can potentially hold different weights corresponding to different types of information, this second type of link-based information can also be subdivided into several levels of cell activation criteria.

### 4.2.2 CONNECTION LINKS

Connection links are generated by connection cells, their primary purpose is to retain the bulk of the data in the system, and their only operation is to find and connect to the most suitable cells. Unlike uniform connection cells, there can potentially be many types of links, depending on the complexity of the data to be represented.

As mentioned these links will complement cells, in that they retain the majority of the data, and perform very limited actions, consisting of only trying to find the most appropriate cell to connect. On the other hand, cells are almost the exact opposite in that very limited data is retained and large amounts of computation comprising data aggregation, decision making and ultimately link generation are performed. Upon the introduction of data into the system, it will be propagated into different c-cells. This operation is facilitated by the connection links, thereby resulting in an *activation path* comprising the c-cells and links currently activated by the particular data.

Similar to c-cells, links can be in one of three involvement states: *dormant*, *inactive*, and *active*. When the link is first generated, or the c-cell to which it is incidental is activated, the link is in the active state. After completing its involvement in immediate operations, its state becomes inactive. If the link is part of a previous activation, and it has still not been activated by the current activation path it is inactive, and can be activated at any time if its incidental c-cell decides that it is suitable as part of its path of operation. This will encourage reuse of previously generated paths, and the overall aggregation and inter-relation of information. It is important to point out that initially the system starts with no predefined links in the most generic mode of the system, and links are generated only as the need arises.

The two principal links required by the system are: intraconnection and hierarchical links. Like their corresponding c-cell operational states, they represent *part to whole* and *type to instance of type*, but unlike c-cells the operational state of each link is fixed upon generation. Essentially, what these links do is transfer information to the c-cells to facilitate their operations, by providing the criteria for decision making and biasing their link generation to follow established paths.

Biologically, connection links would correspond to the dendrites and axons of the neural cell. Different concentrations of ions in dendrites and axons correspond to combinations of weights and operational capabilities of the connection links. The increased number of connection links corresponds to the increased branching of biological links, which is thought to be responsible for higher levels of intelligence. Therefore, the ability to represent complex logical hierarchies is inherently built into the architecture of this system. It is different from symbolic system hierarchical constructs in that both data storage and processing are part of the same structure as in the biological brain and not artificially separated as in traditional SAI.

A very important feature of these links is that they allow certain degrees of adjustability and uncertainty similar to the statistical operation of biological nerve cells. When connecting to a cell, a link may not always find a perfect cell to connect to, and therefore may connect to more than one cell with incomplete degrees of connectivity, it is then the task of the cell to determine if these imperfect connecting links will meet its threshold requirements, and allow the process to continue. Likewise, a generated link may find that a dormant link exists that matches its data requirements, and will choose to simply reactivate that link, instead of establishing a new one. This could also lead to tolerable but imperfect matches that would again be the responsibility of the destination cell to decide upon the links suitability in its internal operations.

### 4.2.3 OPERATION

While the operation of the system is self-enclosed, boot-strapping complementary operations are required to start the system and to access the desired results upon reaching steady state. The central operation of the system is a continuous series of intraconnection and hierarchical operations, referred to as the *core operatives*. Any number of core operative iterations may be performed, with the necessity that the first operation is intraconnection and the last is hierarchical.

## 4.3 SYSTEM VALIDATION

This system from its inception was designed to comply with what is known about biological neural systems and the psychology of intelligence. Its distributed nature and establishment on neural foundations may classify it as a connectionist system, but with the very important distinctions that it is innately capable of representing structure and working with more complex data unlike most connectionist systems. While the type of data held in the system must be carefully chosen, it cannot be considered equivalent to the complex data representations required by symbolic systems. This is because all data stored consists of numerical weights, which are only an extension of the rudimentary weights used in artificial neural networks. Biological neural cells also make use of different concentrations of ions in their data representation, and the many types of dendrites and axons correspond to the different requirements of the perceptual inputs that they interface. In addition, since biological neural cells at the macroscopic level symbolically operate using discrete events in time and space, they can be simulated using numerical data [31]. Therefore, it can be claimed that the system's data representation is only a simplification of that used biologically. The correspondence between a biological nerve cell and that simulated in this system is shown in (Fig. 4.3)

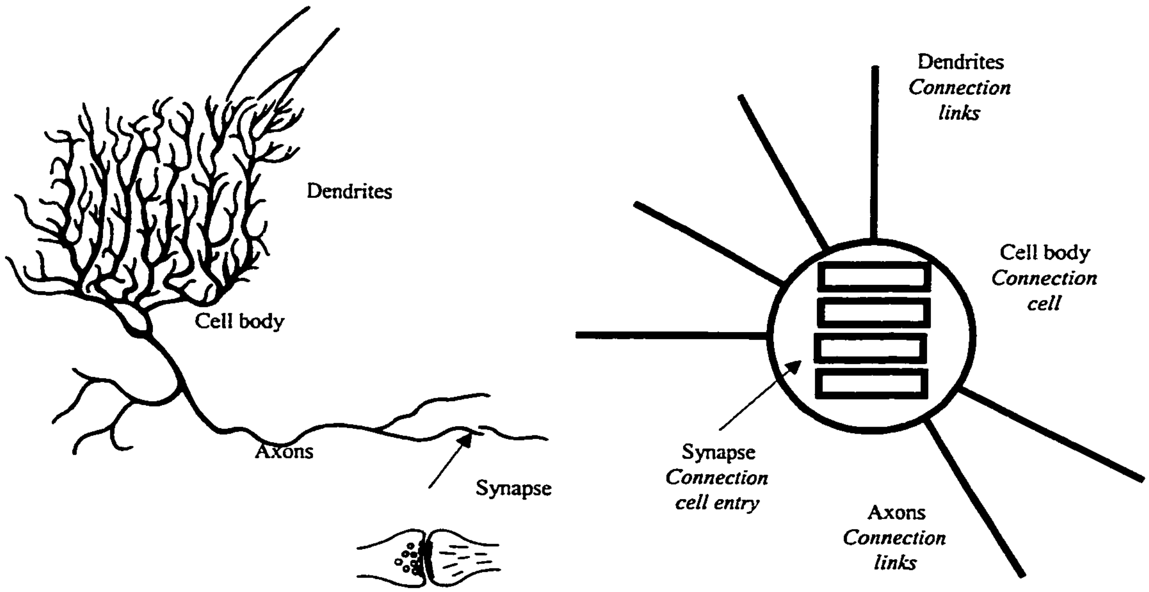


Figure 4.3 Correspondence between biological and simulated nerve cell
Simulated nerve cell consisting of c-cell, links and entries on the right correspond to the generalized biological nerve cell on the left [7].

In direct comparison with current AI architectures, several major differences are observed regarding the solution space (Fig. 4.4). Any architecture must perform a type of pattern matching of the input data into the system and to propagate data through its levels. While this matching is highly constrained for SAI it is extremely flexible and robust for CAI, with the drawback of CAI being incapable of handling multi-dimensional data as in SAI. The CSSC machine allows for both multi-dimensional data similar to SAI and flexible matching similar to CAI. Since the data used in the CSSC machine is not as complex as the data structures used by SAI, matching is slightly less sophisticated. While, the flexibility in matching is not equal to that in CAI due to the existence of structure. These cannot be considered criticisms of the proposed architecture, since the data-structures used by SAI have been shown to be overly complex and the completely unconstrained flexibility displayed by CAI makes its overall functionality extremely limited. As to the connection of multiple solution spaces, SAI provides the ability to generate highly sophisticated hierarchical structures between the different spaces, whereas CAI can manage only rudimentary mapping between its solution spaces. The CSSC machine allows for hierarchical structuring between its different spaces, but at a level that is not as complex as that displayed by SAI. Previously when discussing vision, it was shown that such a high level of sophistication in hierarchy as in SAI is simply not required and is not innately used by humans [16].

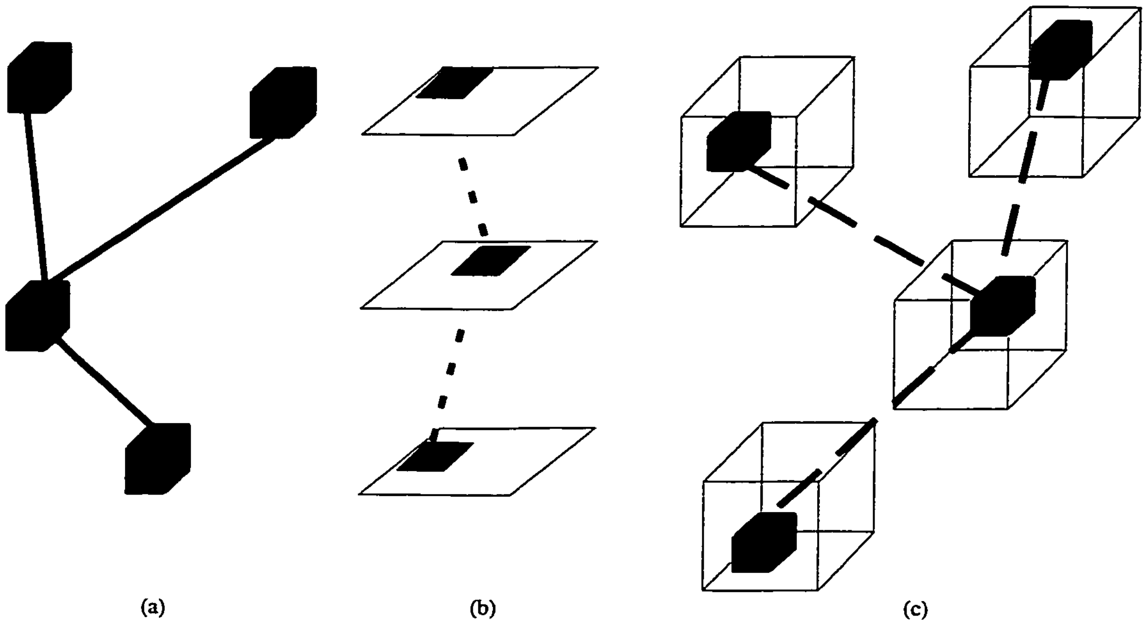

Figure 4.4 Schematic comparison of symbolic, connectionist and CSSC machine solution spaces

(a) Symbolic architectures
Solution spaces allow multi-dimensional matching but are highly constrained, different spaces are connected through flexible complex hierarchical structures.

(b) Multi-layer Connectionist architectures
Solution spaces allow matching in a single dimension but are highly flexible, different spaces are connected through limited unconstrained structures.

(c) CSSC machine architecture
Solution spaces allow multi-dimensional matching that is highly flexible, different spaces are connected through flexible simple hierarchical structures.

It was mentioned that deliberation and sub-deliberation are extremes of a single continuous spectrum of thought. Basic SAI is incapable of sub-deliberation due to its highly constrained solution space and CAI is incapable of deliberation due its lack of structure between solution spaces. Since the CSSC machine includes both flexible solution spaces and the ability to structurally connect solution spaces, it is inherently able to perform both types of thought in addition to any medial modes of thought.

The implemented system is currently capable of bottom-up empirical learning and a limited top-down teaching mechanism provided through a special complementary link. However, no explicit rule forming can be formed at this time. The reason for this was in order not to compromise the design principles in order to achieve a specific task or application that is not in vein with the design of the abstract system. Therefore, the enabling of a top-down learning mechanism would necessarily require the construction of functional components that would be able to directly and homogeneously connect to the knowledge currently stored in the system, and provide a consistent mechanism to interface between the system and the outside world.

## 4.4 CONCLUDING REMARKS

The CSSC machine is an abstract AI framework whose design is based on the complete scope of human intelligence. Therefore, its incorporation of the features of both symbolic and connectionist architectures are not the result of simple hybridization, rather consequential products of the underlying design. The basic goal of this system in its generic state is to organize and relate information in a manner that will make it useful for later complex symbolic functional operations, which are not considered part of the core operatives.

The system's innate ability to self-construct structure based on presented data, is a critical design feature. This feature allows the support of multiple types of data, which is essential for most perceptual processes that operate with numerous types of data. In addition, structure is strongly associated with data itself in terms of both generation and type, this means that the resulting structures are driven by the data rather than a predefined architecture. These two properties of self-structuring are what significantly distinguish this architecture from current symbolic and connectionist architectures respectively.

# CHAPTER FIVE

# IMPLEMENTATION OF CSSC MACHINE FOR COMPUTER VISION

## 5.1 OVERVIEW OF IMPLEMENTATION

In this chapter, the CSSC machine model is used in a computer vision application. The goal of this implementation is to exemplify the use of the model and its core operatives in representing the knowledge required for a sample perceptual problem. As in the cerebral cortex with its distinct functionally optimized regions, complementary application specific operators must be added to the system, in order to best perform the required visual processing application.

## 5.2 VISION AND ITS REQUIREMENTS

The most significant addition to the basic system will be the types of visual data to be stored and operated. An advantage of visual processing is that it is a relatively well-studied area of AI. Therefore, in addition to the goal of discovering the basic elements of vision, the analysis of vision will also help to verify the structure of the basic CSSC machine.

## 5.2.1 THEORY

As discussed in the exposition of vision, all that is required from the image are the cusps and edges to demonstrate primitive visual processing. After these cusps and edges are extracted, they are supplied to the system. The relationship between this initial stage and that of the core system is analogous to the relationship between the visual perception system (optical nerve) and the brain. These representational components are then used to construct the basic shapes that are used for subsequent recognition.

Since this system is to be simulated using a machine that is incapable of performing the massive associability of the optic nerve, extra information needed to be extracted pertaining to the spatial location of each cusp. The lack of associability also presented problems in obtaining the contour of the image. This was solved by using a modified form of chain-coding to provide the sequence of links of the contour, the discovery of areas of deep concavity, and resisting noise [30].

As always, extracting a clean border trace of the input image was a difficult task. Therefore, the complexity of the input images was purposefully minimized, and limited to monochrome images. This is warranted, because it has been found experimentally that the difference in speed of initial recognition of full-color images and their monochrome line-drawing equivalents in humans were almost negligible [14].

## 5.2.2 IMPLEMENTATION

This constitutes the low-level image processing front-end of the system. Its duty is to distill the raw input image into the required data format for insertion into the system. The input image is first monochrome filtered, in order to maximize the probability of clean border extraction. Then a perimeter-tracing algorithm was used for the border detection of the objects in the image [32] (Fig. 5.1). This algorithm first requires the application of a filter to the entire image (Fig. 5.1-a), which embeds into every border pixel of the object the spatial direction of the following border pixel. Once this data is embedded into the image, border paths are followed using the mapping scheme in (Fig. 5.1-b).

This algorithm will detect the exact contour of the object, is highly susceptible to noise and will treat every change in direction of the border as a possible cusp. To deal with this problem, the approximate height and width of each independent object in the image is calculated, then a window of granularity is defined with width and height equal to 10% of the calculated width and height. While tracing the object if a detected cusp is found to be within a window of granularity together with another cusp they are merged into a single cusp, thereby removing undesired noise. Another use of this filtering technique is that it will deal with wire-frame objects that have no hull. Due to the nature of the used perimeter-tracing algorithm a wire-frame object will be initially detected as two adjacent objects (Fig. 5.2-a), the designed merging technique will alleviate this problem. This basic filtering method, effectively removes noise from the object in a way that is adaptive to the size of the object to be analyzed, thereby allowing large and small objects to be diligently filtered and recognized, while not misinterpreting possible features as noise (Fig. 5.2-b).

A second filtering stage is also required to deal with cusps whose two edges approximately constitute a straight line. This is also an undesired type of noise, which may not always be detected using the window of granularity filtering method, particularly for objects that contain perfectly diagonal edges. Therefore, a straight-line filter is applied that will search for all cusps that have an angle of approximately 180° between its two edges and proceed to remove the cusp and join the two edges into a single edge.

A final filtering stage is also required to merge cusps that are not along the same path, but still within the window of granularity. Due to the inability of digital images to display perfect diagonal lines, recognize "wire" shapes (Fig. 5.2-a) and over-emphasize noise, these three filtering stages are required. Their particular order was found empirically to give the most reliable results over a variety of different images.

Once the cusps have been established, the lengths of the edges must be discovered. In order to simplify the process and maintain acceptable performance, it was decided that all edges are straight lines, with no curvature. In addition, it was assumed that no more than two edges can be incidental to a single cusp, again this was in the interest of simplification and minimization of data. The straightening of edges can also be regarded as another stage of noise removal in order to further simplify analysis. The length of an edge is found by a simple application of the distance formula between the two cusps that bridge the edge. However, not all cusps will be connected to two edges, as in the case of an end-point cusp. These end-point cusps were previously determined through window of granularity filtering and subsequently marked as end-points. Since the number of edges per cusp is limited to two, this means that in any connected object there will be two end-points or none at all. By connecting the two endpoints with a *pseudo-edge*, the biological ability of edge completion is satisfied. While this pseudo-link is later discarded before entering the system, it is first used in providing additional directional information as to be explained.

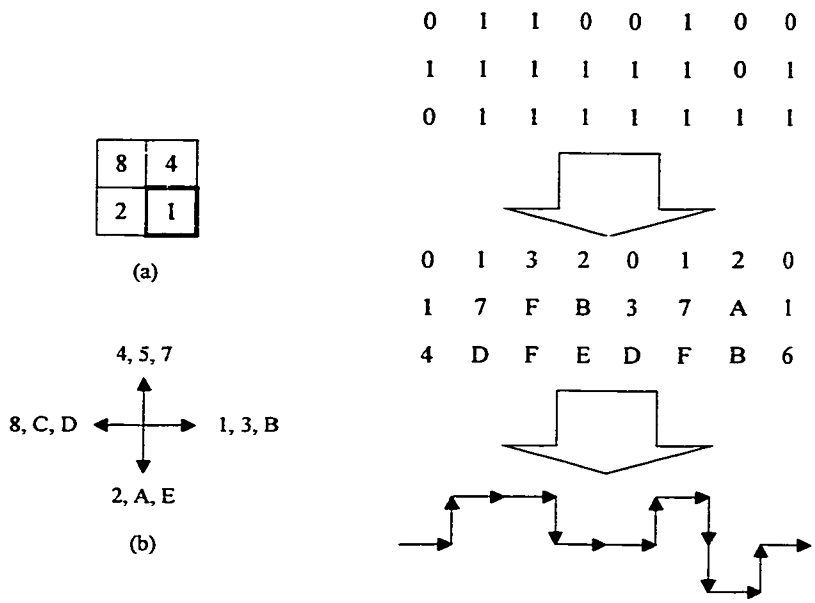


Figure 5.1 Perimeter tracing algorithm
Filter in (a) is applied to the entire image, then the resulting image is traced as in (b) according to its hexadecimal value.

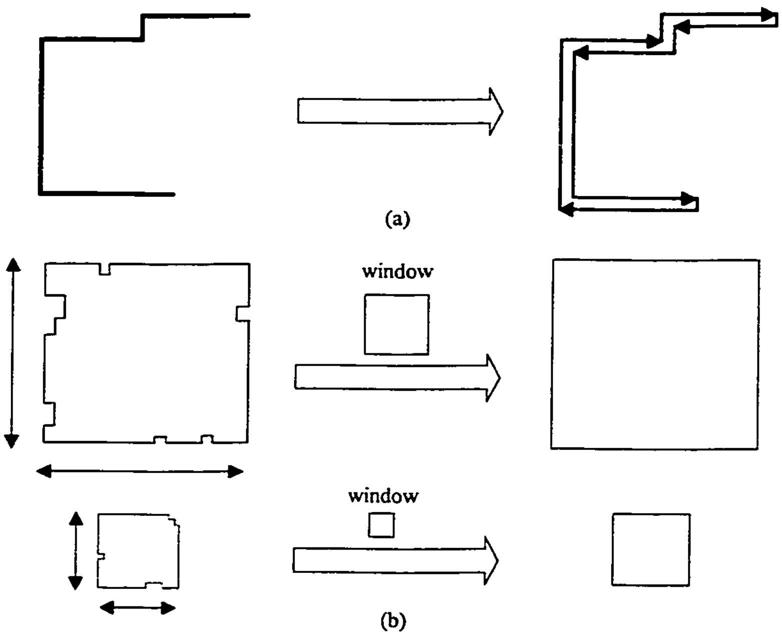


Figure 5.2 Window filtration scheme
(a) Wire-frame objects will be initially traced as two adjacent objects.
(b) Adaptive window sizes will help to appropriately detect noise in different size objects

In order to maintain a sense of spatial associability between the components of an object or even the entire image, a method of association must be established. It was decided that the angular direction of edges would determine this spatiality as a substitute for the retina's associative ability. While at first it may seem that the use of direction is too discrete compared to its biological original, many steps were taken to reduce its rigidity. First, no exact measurements are expected in any of the systems operations, and this is done through the use of a suitable match-tolerance. Second, all angles are constrained to be between 0° and 90°, this adds an element of rotational invariance, further moderates the measured values and simplifies the descriptive terminology to only "next-to". The actual calculation is done for every edge by obtaining the slope and using it to determine the angle in degrees (Fig. 5.3).

Once, this low-level image processing stage is complete, the pertinent data of the image is ready for insertion into the system. Each data entry consists of the spatial location of the cusp, the angular direction toward its two neighbors, and the length of its one or two edges. Additional data such as the order of the cusp during object traversal is discarded, since its usefulness in determining what object the cusp belongs to is now over.

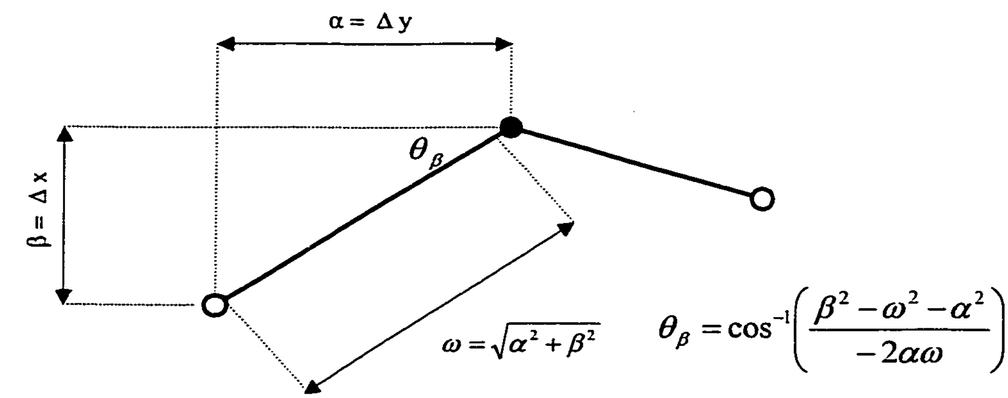


Figure 5.3 Calculation of edge direction
Distance-slope formula is applied to every edge.

## 5.3 DESCRIPTION OF IMPLEMENTATION

The following description will explain in detail how the core system is used along with complementary operators in order to achieve a basic visual processing system.

### 5.3.1 CONNECTION LINKS

Connection links are the main data retention components of the system. Their task is to distribute required data to c-cells for processing, and to add structure to the system by establishing stable connections between the c-cells. Their basic mode of operation in this system is shown in (Fig. 5.4).

#### 5.3.1.1 CORE LINKS

Intraconnection links (i-links) and hierarchical links (h-links) compose the main knowledge representation links in this system. Together i-links and h-links are able to relate and group objects together and to establish an adequate hierarchical structure. Upon introduction of an image to analyze and the availability of data in suitable form, these two links will operate sequentially until the image has been completely analyzed.

While these two links denote the core links of the system and consequently perform the bulk of the processing workload, they still require supplementary links to deliver data in the required form, extract the processed information, and relate to the external user. However, in any implementation of this system, these same two core links

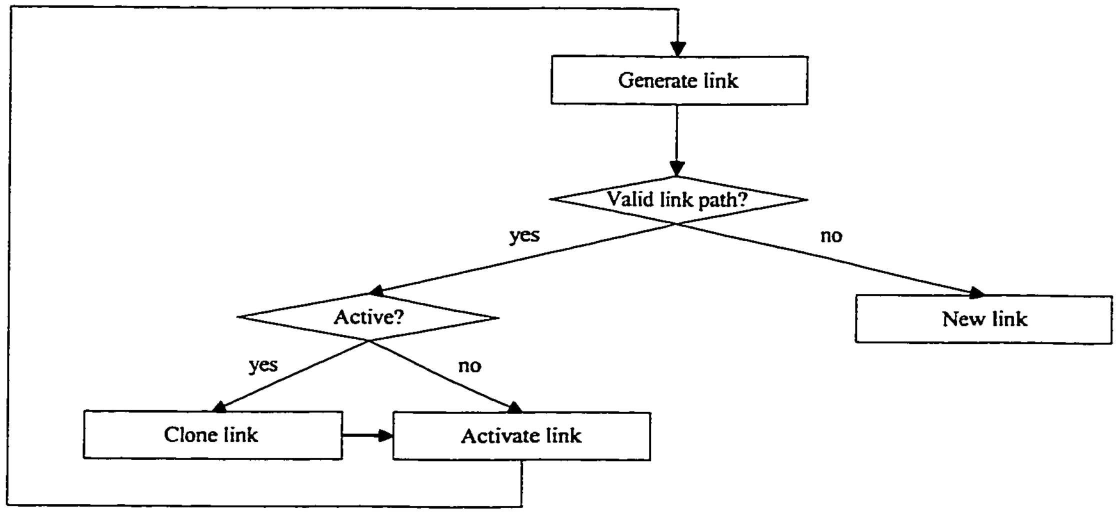


Figure 5.4 Operation of connection link generation

would be used. The only significant difference would be in the number and complexity of the supplementary links.

In the discussion below, i-links and h-links will be discussed for the particular application of computer vision, with particular emphasis on how they process visual information, and how this processing closely corresponds to the identified operations of the biological visual system.

5.3.1.1.1 Intraconnection Links. These links represent the part-to-whole structure and in this context, they connect different components of an *object*, where an object is any visual entity that can be analyzed alone. For the instance of a single connected object, such as a circle or square, the i-links would represent the connections between the different components that make up the object. If the object consists of a collection of non-incidental simple objects, the i-links would relate the spatial relationship between the components making up the object. An example of this would be a simple face that consists of two simple eyes a nose and mouth, where each component is regarded as a simple object and the particular spatial relationship between them would construct a face. The rationale behind this similarity in treating simple and complex objects, is to have a seamless intraconnection mechanism that will effectively deal with as may different visual scenarios as possible using only a single abstract mechanism.

Intraconnection links when relating two components use the normalized angle and distance to quantify their relationship and add a sense of associability between them. Angles are used to relate the objects spatially, and as mentioned, the use of normalized angles partially increases the degree of rotational invariance and reduces the discreteness associated with the use of complete angles. Distances between objects are also stored because they provide additional spatial information between components that is useful in later decision-making stages.

Components that need to be related with i-links, may or may not have been previously related. First, the new components' spatial relationship is calculated. If i-links already exist between the components and they have approximately the same angular relationships, the i-links are considered a match. Now, these matched i-links may be currently active or inactive. If they are inactive, the i-links are simply activated and their spatial information is adjusted with the information of the new components. If the i-links are already active, they are cloned and the adjustment process continues as in the inactive i-link case. Supposing that the components have not been previously related, or that the existing i-links do not closely match with the spatial information of the new components, new i-links are generated with the spatial information of the new components. This component intraconnecting procedure is the same for simple objects where the components are the edges of the contour, for complex objects where the components are single objects, and for entire images where the components are complex objects. As explained, previously existing i-links can bias the linking of components and make use of previous experience. When new i-links are generated, it is part of learning a new experience. When an i-link follows a previously established path, it is considered a one-shot recognition, whereas the partial or complete generation of new paths corresponds to the application of a relaxation algorithm. The actual strength of this technique, is displayed with partial activations, that is, relating of components using old and new i-links. With partial activations, previous experience can help in recognizing unfamiliar objects by recognizing certain components and aiding in the overall recognition effort.

An important element of intraconnection is that it will try to relate all present components together. While this is not a problem with simple components, since only incidental components are related, and the relations between non-incidental components are automatically deleted. A problem may exist with complex objects, if they are not progressively introduced to the system. For example, if the systems first introduction to an image of a face, is that of a face with a hat, it will consider this as an exemplar of the category. This would lead to the later recognition of a hat-less face, as being an aberrant exemplar of the category. It is important to recall that this system, which attempts to simulate intelligence will not be able to focus its attention to certain objects, while neglecting others. This capability to focus attention is a component of mind, which encompasses intelligence among other components. Therefore, for optimal performance this current system must be gradually introduced to different objects, in order not to corrupt its categorization abilities.

5.3.1.1.2 Hierarchical Links. This link represents the type to instances of type hierarchical structure, and is the primary source of information grouping and decision-making for the c-cell. Following the intraconnection of components of an object, h-links are generated from the source c-cell of each component and are grouped into a c-cell entry. For example, if the components of a simple object such as a circle have been intraconnected, the involved c-cells would generate h-links that are clustered together into a single c-cell entry. This would mean that if the same components were later activated and subsequently the corresponding h-links and the c-cell entry were activated, then the particular activation is also “circle”. Similarly, in a more complex situation example such as that of a face, each of the face’s components would first be clustered into types, such as eyes, nose, and mouth. Then following intraconnection of these components, they would also be clustered together into a single entry representing the entry “face”. As in i-links, an attempt was made to keep the h-link process abstract and as robust as possible to deal with as many visual scenarios as possible.

H-links retain only distance information provided to them by the previous intraconnection stage. This distance information together with the total number of physical h-links facilitates the categorization of the object. . The number of h-links in the activation group is a coarse type of data, while the individual distances are finer forms of data. Both data types are used in concert in the decision making process of the c-cell.

When a pattern of intraconnection has been activated and h-links are to be generated a number of different scenarios may occur. The default case, of no previously existing h-links propagating from the activated c-cells results in the generation of a single h-link per c-cell retaining the distance information from the intraconnections, into a new c-cell entry and is considered an exemplar of an unfamiliar category. In the perfect match case, there exists a single h-link per c-cell and they are all clustered into a single entry that is satisfactorily activated, then this particular activation is considered an exemplar of a known category. If the c-cell entry was not satisfactorily activated, then the path is considered false and new h-links are generated to cluster into a new c-cell entry as in the default case. The default and perfect cases are operational extremes, what usually happens is that in a single activation some c-cells have no propagating h-links, some have a single propagating c-cell, while others have more than one propagating h-link. The solution to this typical scenario is to activate all incidental h-links, and examine the corresponding c-cell activations. If only a single c-cell entry was activated then only that particular h-link path is activated, if no c-cell entries were satisfactorily activated then the default operation of new h-links and c-cell entry is pursued. However, if more than one c-cell entry is activated, the object is considered a possible exemplar of both types and the correction of this activation is done by selecting the path with the closest match as the desired activation. This multi-activation situation occurs frequently with an inexperienced system, a simple example is that of confusing squares and rectangles. This confusion is later rectified with increased experience.

It is worth noting that h-links do not store or make use of the angle data held by i-links. This is because a decision based on the angular data was already done explicitly by the i-links upon intraconnection and does not need to be repeated again. This is another example of the important design characteristic of this system where decision-making is localized at every stage and is not repeated. This separation of object recognition into two stages that use dissimilar information in accomplishing their task, also allows much more sophisticated categorization than that of traditional connectionist systems and an element of multi-dimensional matching.

5.3.1.2 SUPPLEMENTARY LINKS

The following two links: visual links (v-links) and external links (e-links) are unique to this vision system in that they are primarily designed to help interface the system and its core operations to the outside world. The v-links and e-links affect the systems operation only in the first stage to interface with the operational environment or the last stage to interface with the mentor respectively.

5.3.1.2.1 Visual Links. These are the only links of the system that are not generated by c-cells, rather they are directly generated upon viewing an image for analysis. The role of these links is to boot-strap the system, by obtaining the recently extracted visual data and placing it in the most suitable c-cells.

As mentioned previously, the system considers the bulk of the visual data to be stored in the cusps, and the remainder in the connecting edges. A single v-link is generated per cusp, it holds the spatial location of the cusp, and the distance and direction of its incidental edges.

Usually a new v-link is generated for every cusp. However, if an inactive v-link already exists and maintains data that is satisfactorily similar to the new data, it is activated and the data is adjusted. This reactivation process usually only occurs with very similar data, since the matching criteria is quite stringent for both angles and directions. The assignment of v-links to c-cells is designed to place symmetrically corresponding cusps into a single c-cell, this correspondence is calculated using both angles and distances of the cusps edges. In the case of cusps not having symmetrical correspondents, they are assigned to c-cells that previously contain v-links with similar directional values. If the cusp does not fit into any of the previous assignment categories, it is assigned to a new c-cell. This assignment technique facilitates the activation of the correct pattern and aids further stages of the system.

V-links act as the optical nerve of the system, providing it with only the pertinent data, and filtering out any data that is not required. Its only task is to start the visual process and interface between the low-level image data acquisition component and the main system.

5.3.1.2.2 External Links. These links interface the system to the external mentor. After the system has reached its steady state and all the components of the object have been clustered into a single c-cell entry, e-links are invoked. The first task is to announce the findings of the analysis, either the object has been recognized as an exemplar of a single known category or multiple categories, or as an exemplar of an unknown category. The second task is to translate the mentor's reaction to the systems findings back into the system, and perform any rectifications.

As mentioned, each c-cell with h-link entries is capable of retaining a unique label. If any of its h-link entries are activated and the process has reached steady-state, an e-link is generated to announce the contents of this label to the mentor. Based on the reaction of the mentor, different actions are taken by the e-links. Therefore, it can be seen that e-links are not physical like the other links that hold particular data, they are in fact only abstract mechanisms used to relay information between the system and outside world.

Depending on the reaction of the mentor, different actions are taken by the e-link. If the activated entry belongs to a c-cell with no label, the e-link announces that it has found an exemplar of a new category. The mentor can then either assign the category with a label assuming that it is a significant category, or leave the label blank if the category is considered insignificant to warrant a label. The e-link will act according to this assignment, in case the label has not been assigned previously to a c-cell, it is assigned to the new c-cell and the activation pattern is considered a perfect exemplar of the new category. If the label has been assigned previously to c-cell, the activation pattern is considered an alternative exemplar of the existing category and its h-links and c-cell entry are assigned to the correct c-cell, this case may occur when viewing objects from different aspects. If the activated entry belongs to a c-cell with a label, the e-link announces that it has found an exemplar of a known category. If the mentor confirms the e-links findings, the h-links weights are adjusted to compensate for any missing components. This compensation better characterizes the importance of different components, since by default all h-links are given equal weight. If the mentor refutes the findings, it means that the pattern of activation is not an exemplar of that category. The subsequent action depends on the nature of the label. If the label already exists, the activation pattern is considered an alternative exemplar of the category, new h-links and c-cell entry are generated and assigned to the correct c-cell, and the incorrectly activated c-cell entry decreases it tolerance to curb future mistakes. If the label does not exist, a new c-cell is chosen to carry the label, new h-links and c-cell entry are generated and assigned to the new c-cell, and the incorrectly activated c-cell entry decreases its tolerance.

These links play the role of the top-down teacher, with them incorrect conclusions reached by the system can be rectified and correct conclusions can be confirmed. An important capability is the capacity to assign multiple objects, which may initially seem totally unrelated to the system, a single label. This may happen in the case of multiple views of a single object, or multiple combinations of objects that constitute a single label. Of special importance is the method of teaching, particularly if it is desired to construct a hierarchy of categories and sub-categories. Therefore, it is vitally important to train the system in a regulated manner in order not to disrupt any previously learned information or limit its future learning capabilities.

### 5.3.2 CONNECTION CELLS

Connection cells (c-cells) are the main information processors in the system. They receive input data from incoming links, and their output decision consists of generating links or not. C-cells for this system are all homogeneous, that is they all have the same processing capabilities initially, but their responsibilities toward the system's operations change as information is stored, processed and links are generated.

All entries contain certain common information whose main goal is to facilitate associability between the components of the simulated system. The first common data is the degree of activity, which ranges from *dormant* meaning that a link was previously established but is not currently active, to *active* meaning that a currently connected link is active and part of the current pattern of activation. An additional degree of activity, *inactive* was also introduced as an intermediary degree of activity between dormant and active. The reason for this intermediary state is due to the serial nature of the simulation, which meant that only a single processing path could be active at a time. An example of this is an image consisting of two components being analyzed, first, one component is analyzed that becomes inactive, second the other component is analyzed, and finally both components become active for the final merging into a single object. The second common data, is spatial location of the object being analyzed, this is also necessary for the simulation in order to help with the associability of the components in a serial processing environment.

When a link connects to a c-cell, it places an entry within the c-cell that establishes connectivity to the link and facilitates the processing and subsequent transfer of information to successive links (Fig. 5.5). Biologically these c-cell entries can be likened to synapses, in that they are dynamically generated and are the main point of information gathering and processing.

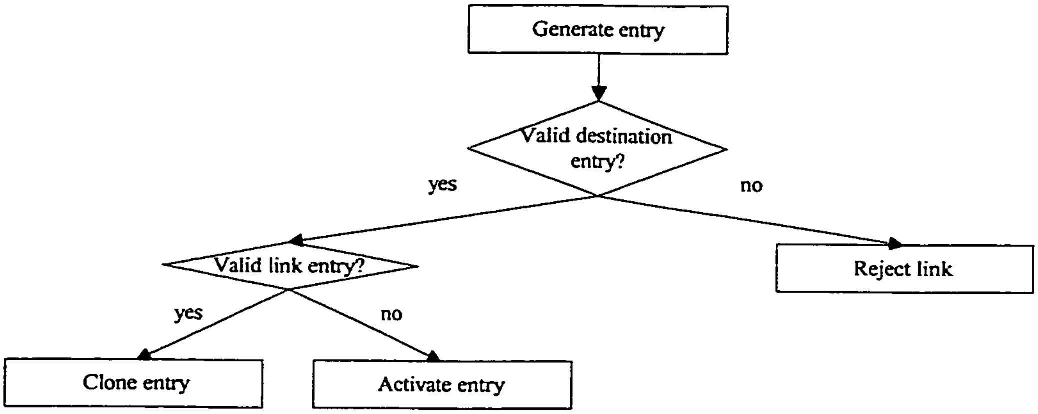


Figure 5.5 Operation of connection cell entry generation

### 5.3.2.1 LOCAL INFORMATION

The c-cell needs to hold some local information, which is not data that can be processed, rather it will help monitor the ongoing activities of the c-cell itself. Since a c-cell fluctuates between operational states depending on the type of incoming active links, it is necessary to assign a state variable to each c-cell. If the c-cell is dormant, meaning it has no incoming active links or it has reached steady-state, it is assigned *state-0*. In case it has incoming active v-links or h-links, its state becomes *state-1* and becomes ready to generate i-links, this is also referred to as the *growth state*. If the c-cell is in *state-2* the *propagation state* it means that it has been intraconnected with active i-links and it is ready to generate h-links (Fig. 4.2).

If the c-cell has h-link entries, it means that it has been assigned a class and will therefore have a label. This label is accessed only by e-links and its only use is to facilitate interaction with the mentor. It is not an intrinsic system data type, but its replacement would require the establishment of an entire language processing system that would be accessed by this visual processing system.

### 5.3.2.2 VISUAL LINK ENTRIES

These entries are established by active v-links that connect to a dormant c-cell. They are responsible for changing the state of the c-cell to state-1 and activating the entire system. The entries are copies of the information passed by the links, namely the cusp spatial location and the directions and sizes of up to two corresponding edges.

Upon c-cell activation, the c-cell will generate i-links to other state-1 c-cells that contain v-link entries that correspond to its own. It does this by searching for an active v-link entry that has a common edge, with the same size and inverted angle, which meets the c-cells matching threshold. This matching technique could result in some problems when more than two components have common edges. That is why symmetrical components with similar size edges and inverted angles are assigned to the same c-cell in order to alleviate this possibility. Therefore, with the current specifications of the system, each v-link entry can generate up to two i-links, consistent with the number of edges per cusp. At the end of this process, all the components of the object should be intraconnected.

It can be seen that the v-link entry is only a copy of the v-link data localized to the c-cell. Once i-links are generated and the data is propagated, the data is of no use to the c-cell and can be purged since the information is retained by the v-links.

#### 5.3.2.3 INTRACONNECTION LINK ENTRIES

Upon intraconnection of components, an entry is established for each i-link connecting to a c-cell. The c-cell changes its state from state-1 to state-2, is ready for the generation of h-links, and upon completion will return to the dormant state-0. Each i-link entry will carry the distance and angle information for a single edge.

Once all components of the object have been intraconnected and the states of all the involved c-cells have changed to state-2, every involved c-cell will generate h-links. If only a single i-link entry is in the c-cell, the generated h-link will represent only half a component since only one edge is involved. If there are two i-link entries in the same c-cell that correspond to the same cusp they are merged together and the generated h-link will represent a complete component. If the c-cell contains i-link entries from more than one cusp, more than one h-link is generated according to the previous two cases. This process demonstrates the importance of correct initial placement by the v-links, since incorrect placement would result in a completely different pattern of activation.

Intraconnection involves all state-1 c-cells in the current activation process. Therefore, the source of these components may be from v-link or h-link entries with no difference in the data format or process sequence of the i-link entries.

### 5.3.2.4 HIERARCHICAL LINK ENTRIES

H-link entries are the result of clustering h-links generated from the intraconnection of a single object. If the h-link entry is activated it will change the state of the c-cell from state-2 to state-1 and intraconnect with other state-1 c-cells if more than one state-1 c-cell exists, or if it is the only remaining active c-cell the system is considered to have reached steady-state and an e-link will be generated. Data in the h-link entry consists of the total number and cumulative weights of the active connected h-links, and the allowable tolerances of activation.

When an h-link entry is first generated, the particular pattern of activation is considered a perfect exemplar of that category and the number of h-links and their cumulative weights are stored as optimal activation parameters. Initially, the activation tolerances are set at half the number of h-links and total weight. Upon subsequent patterns of activations, two types of entry activation may occur. Activation can be a *perfect match* meaning that the activation pattern is satisfactorily close to the exemplar activation pattern in terms of the number of h-links and their weights. It is satisfied if the following two conditions are met:

1. Number of currently activated h-links = Number of exemplary activated h-links
2. The following inequality must be satisfied for every h-link weight:

$$\alpha - \left[\frac{(1 - tolerance) \cdot \alpha}{2}\right] < \chi < \alpha + \left[\frac{(1 - tolerance) \cdot \alpha}{2}\right]$$

| | |
|---|---|
| $\alpha$ : | Exemplar activation pattern value |
| $\chi$ : | Current activation pattern value |
| *tolerance* : | H-link entry weight tolerance |

The other type of activation is a *scaled match*, this may mean that the activation pattern is a scaled version of the exemplar activation pattern with different edge lengths, or that the pattern of activation has an allowable number of missing components, which were originally present in the exemplar activation. Scaling of an image is performed by multiplying each point of the original image by a scaling matrix:

$$[x', y'] = [S \cdot x, S \cdot y] = [x, y] \cdot \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix}$$

$x, y$ : Original image coordinates

$x', y'$ : Scaled image coordinates

S : Scaling factor

Therefore, scaling activation is satisfied if the scaling factor is equal for all existing edges of the object:

1. Number of currently activated h-links >

   (Number of exemplary activated h-links $\cdot$ *tolerance*)

2. The following inequality must be satisfied for every active h-link weight:

$$ratio - \left[ \frac{(1 - tolerance) \cdot ratio}{100} \right] < \chi < ratio + \left[ \frac{(1 - tolerance) \cdot ratio}{100} \right]$$

$$ratio = \frac{(ratio + \chi)}{2}$$

*ratio*: Average activation pattern value, initialized with the first activated pattern value.

$\chi$ : Current activation pattern

*tolerance* : H-link entry number tolerance in (1) and weight tolerance in (2).

If the h-link entry is activated and the result is later refuted by the mentor, the tolerances of the entry are adjusted according to the type of its activation. If the activation was perfect, it means that with the exact number of required h-links the entry was activated. To fix this erroneous activation, the total weight tolerance of the entry must be raised, to avoid similarly incorrect activations.

$$tolerance = 1 - \frac{\text{Total received weight} - \text{Total required weight}}{\text{Total required weight}}$$

For incorrect scale activations, the tolerance of the h-link entry for number of activated links must be raised. This is in order not to repeat similarly incorrect activations, which might have been the result of not considering enough active h-links.

$$tolerance = \frac{\text{Number of currently received active h-links}}{\text{Number of exemplar active h-links}}$$

It may also happen that the pattern of activation resulted in a correct scale activation of the h-link entry, with the number of active h-links less than the exemplar number of h-links. This means that the active h-links can be considered more important than the other inactive h-links, and should have their assigned weights strengthened to reflect their greater importance in the classification of the object at the cost of weakening the inactive h-links.

$$strengthen = \frac{off \cdot on^2}{(off + on)^2} \qquad weaken = \frac{off^2 \cdot on}{(off + on)^2}$$

| | |
|---|---|
| *off*: | Total number of incidental inactive h-links |
| *on*: | Total number of incidental active h-links |
| *strengthen*: | Amount added to the weights of each active h-link |
| *weaken*: | Amount subtracted from t he weights of each inactive h-link |

When h-link entries are to be intraconnected as in the case of an object that consists of several component objects, the generated i-links treat the medial location of the object as the component cusp and the distances and angles between different objects as edges. This seamless transition between components, objects and entire images reduces the complexity of the systems operations.

## 5.3.3 OPERATION

The system is activated upon the introduction of a new image for analysis. The new image is first analyzed by the low-level visual processors to remove noise and extract the pertinent data into the form required by the system. V-links will transfer this extracted data into the core system by attaching to the most suitable c-cells and creating v-link entries to temporarily hold the data. Once v-link entries have been created the systems core processes are started.

The core system processes always start with the Intraconnection of the v-link entries using i-links. Following intraconnection, i-link entries are established in the destination c-cells in order to propagate required data to the next stage. The second stage follows the intraconnection of the components and clusters them using h-links into a single class h-link entry.

If there is only a single remaining active c-cell it means that the system has reached steady state and is ready for interaction with the mentor using e-links. However, if more than one c-cell is active, it means that the image consisted of more than one object and the process continues with the intraconnection of these new components and their clustering into a single object. If a steady-state is still not reached the process is continued.

## 5.4 ANALYSIS

This system was designed to simulate a primitive visual processor using what is known about the human visual processing system biologically, physiologically, and psychologically. Therefore, while some aspects of the system's operations have not been discussed and corroborated by the relevant literature, to my knowledge none of the operations contradict with any conclusive experimental results published thus far.

This demonstration vision system displays the utility of the basic intraconnect and hierarchical operations of the basic system. With the addition of only simple complementary front-end and back-end mechanisms to interact with the required application, a rudimentary vision processing system was constructed. If the complementary mechanisms were modified to extract more information such as edge curvature, image depth, or color for the low-level component and v-links, a more faithful interpretation of the biological visual system can be constructed. The back-end e-links as mentioned are only abstract mechanisms in place of a language processing system, with its addition not only would this system be greatly enhanced, but much more complex visual-linguistic applications may be possible.

# CHAPTER SIX

# EXPERIMENTAL RESULTS

## 6.1 OVERVIEW

In this chapter, the developed computer vision system is tested on a variety of images. While the test images are simple, each is designed to individually demonstrate a single capability of the system. Images of higher complexity would use these same basic capabilities while being entered into the system and subsequently recognized.

## 6.2 LOW-LEVEL COMPLEXITY DEMONSTRATION

The goal of this experiment is to demonstrate the system's ability when applied to simple single component objects. While it will not demonstrate higher levels of structure, it will effectively show how the system is able to learn and integrate simple objects into its memory.

### 6.2.1 EXPERIMENTAL SETUP

The system starts with no previously learned objects and is introduced sequentially to five different objects (Fig. 6.1). Every object has a corresponding figure highlighting its modifications (Fig. 6.2-6.6), the upper part of the figure is of the system

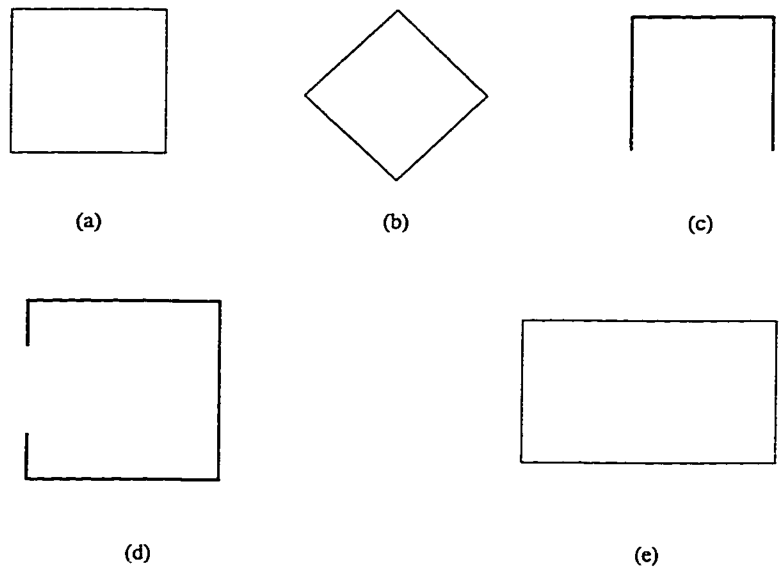


Figure 6.1 Objects involved in low-level compexity demonstration
Objects (a)-(e) were analyzed sequentially by the system, they have been slightly enlarged for this figure but are still to scale.

being introduced to the new image, the middle part is of i-link generation, and the bottom part is of h-link generation. A path with dashed lines conveys a previously learned but dormant path and a path with solid lines and transparent interior is an inactive path being used by the current image. If the path is involved in the current image and is active, it either, has a solid interior meaning that it is completely activated or a hashed interior meaning that it is partially activated.

### 6.2.2 EXPERIMENTAL RESULTS

The goal of this set of experiments is to demonstrate the systems ability in dealing with many of the visual scenarios it might encounter. Representative figures follow the explanation of each object's operations.

#### 6.2.2.1 OBJECT ONE

This is a simple square with right angles. Upon introduction to the system, it is first analyzed by the low-level image processor and found to have four cusps each with two edges. They are identical except for the order of their edge direction angles, as such, a line of symmetry is found and the two cusps are assigned to two different c-cells.

The four cusps are intraconnected with two i-links per cusp connecting to the neighboring cusps, resulting in a total of eight i-links and h-link entries. Following this growth phase, the involved c-cells start the propagation phase. A single h-link entry for each cusp is created, with the corresponding information from its two edges. The four h-links converge into a single h-link entry and since the entry is new e-link conformation assigns it the label "square" (Fig. 6.2). It can now be said that the system has learned what a "square" looks like.

<u>6.2.2.2 OBJECT TWO</u>

This object is also a square but its overall size is slightly smaller than the previous object and it is tilted diagonally at a 45° angle. The four cusps each have two edges, but with identical direction angles for each edge. Since the edge angles and sizes are not similar to the previous objects parameters, the existing v-links cannot be reused. However, since the angles between the edges are right angles and equal to those of the previous object they are to be assigned to the same previously used c-cell, thereby relying on a previously learned path of activation to initially guide its entry into the system. This object has multiple lines of symmetry, therefore all the v-links are assigned to a single c-cell.

Intraconnection is done between the v-link entries in a single c-cell. Previous i-links cannot be reused because they are between two c-cells, which are not both activated by this object and subsequently the associated h-links are not activated. New h-links are generated from this single c-cell in the same manner as the previous object of one h-link per cusp. The h-links converge to a new h-link entry and are assigned the label "diamond" by the mentor (Fig. 6.3).

This object is largely similar to the previous object, except for its diagonal rotation. Therefore, the system could have been designed such that the two objects would be assigned to the same activation path. However, in the interest of biological conformity complete rotational invariance is not allowed and only rotational resistance is allowed up to a small tolerance of approximately 15°. While this tolerance is small compared to that reported in humans, it is necessary to keep this value low due to the coarse granularity of the digital images and the approximation of edges as straight lines. This implementation of rotational invariance is in line with the argument presented in [11]. The argument stated that physical position provides plasticity rather than the visual system. In the current system, since physical position is fixed, all images are read in the same context and it is then the task of the inner system to attempt to recognize the image in its most familiar stored rotation. This is analogous to reading upside down, where the letters and words are easily recognized, but the slow down in reading speed can be attributed to internally rotating the image to find its best match.

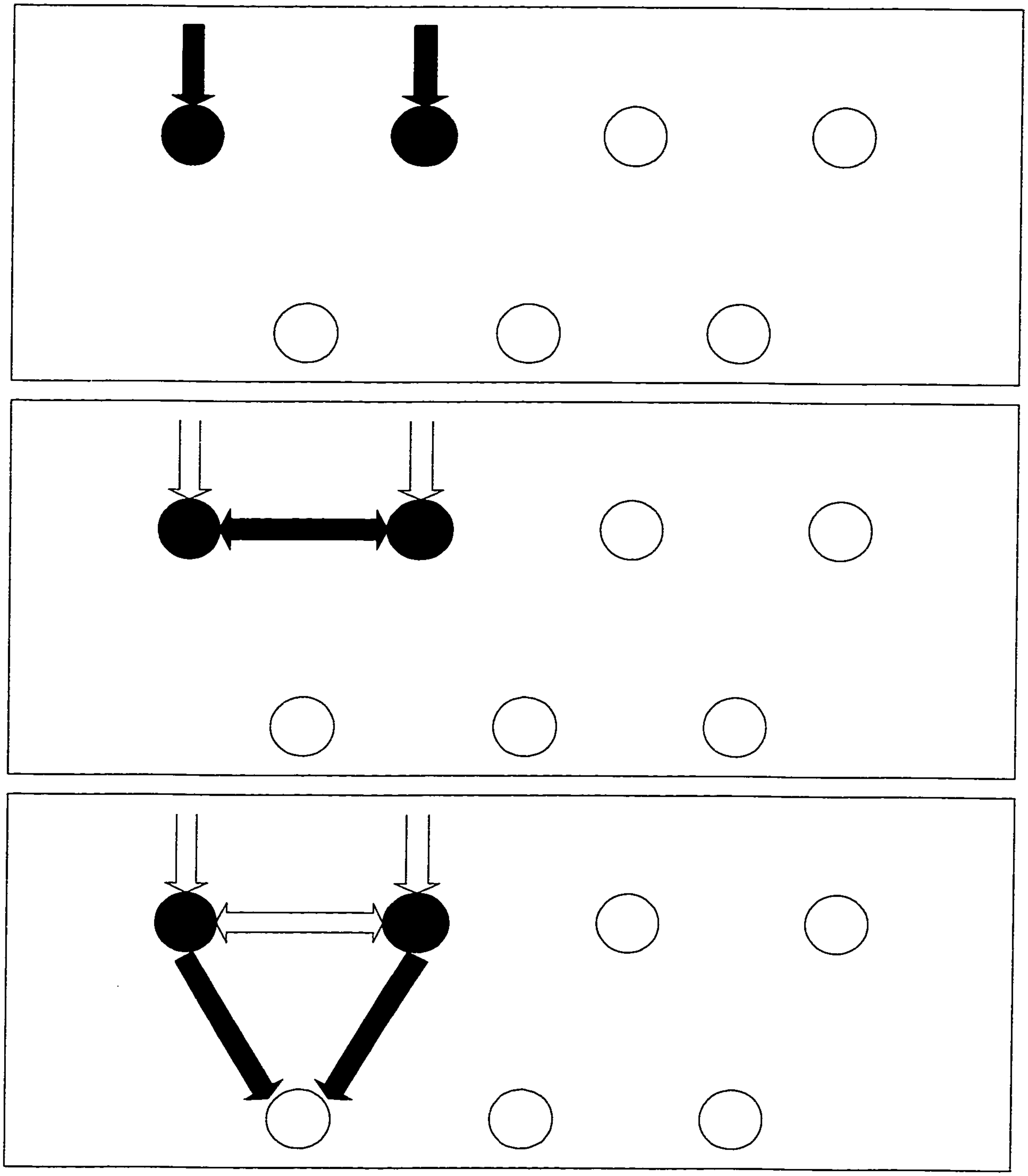

Figure 6.2 Demonstration object one
This is a new system being introduced to the first simple object.

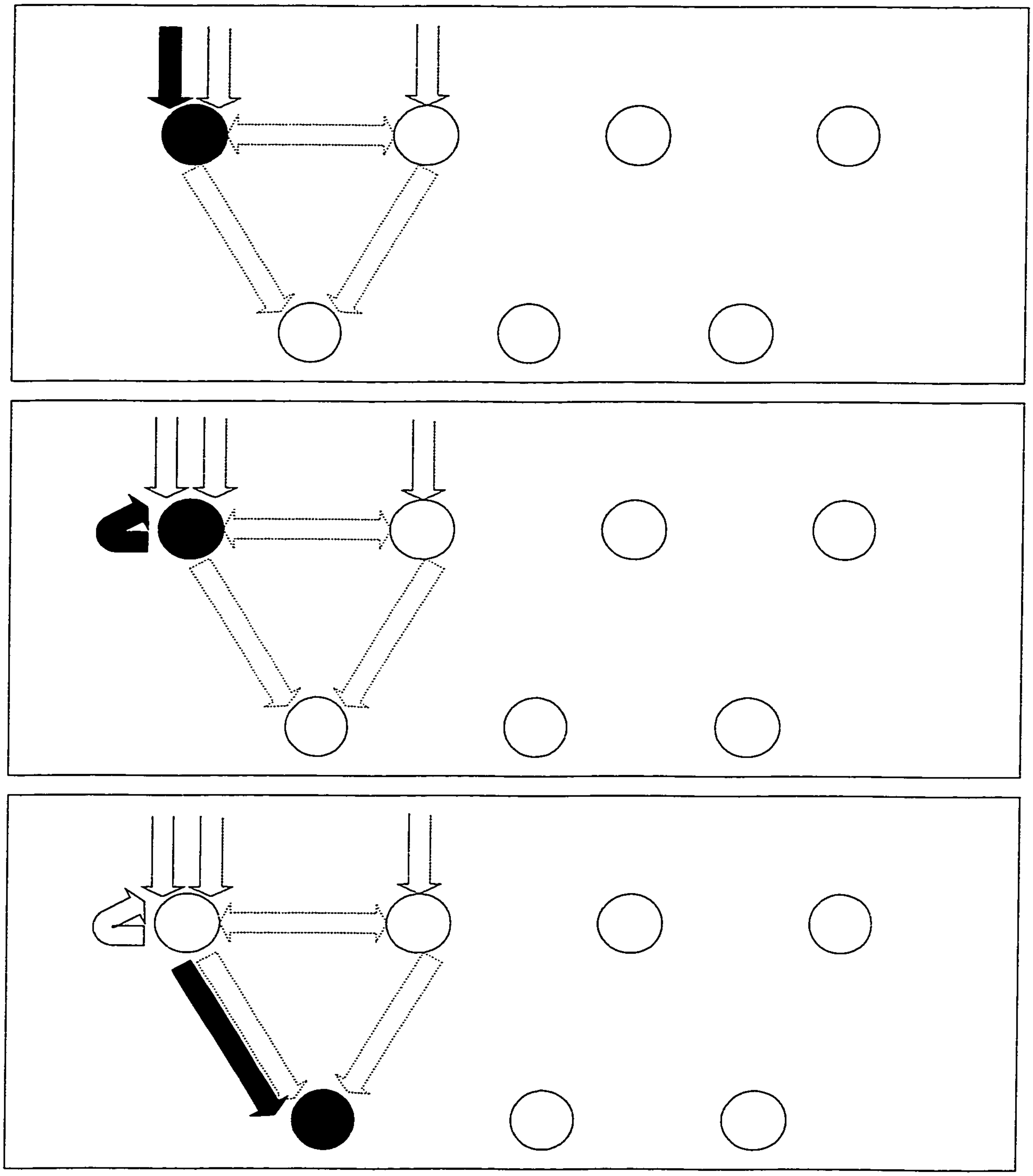

Figure 6.3 Demonstration object two
Previous knowledge will influence the placement of the v-links, and subsequently influence the placement of the entire activation pattern.

### 6.2.2.3 OBJECT THREE

This object is an exact copy of the "square" object except for a missing edge. Upon initial analysis two cusps are extracted each with two edges and two cusps having only a single edge and a pseudo edge each. The pseudo edges are an approximation of the distance between the two single edge cusps and are used only for determining the angles of the other neighboring edges. These cusps activate the same v-links of object one, except for assigning a null distance value in place of the missing edges.

Intraconnection also makes use of the previous i-links, since both c-cells are activated. However, the object's missing edge will correspond to the activation of only six i-links, because every edge is shared between two cusps. With the i-link path activated, the corresponding h-link path is also chosen for activation. Two h-links will have the same activation values as object one, while two h-links will have half the activation values because of the missing i-links. This will result in the scaled activation of the h-link entry and recognition as "square". Upon confirmation, the weights of the two h-links with missing i-links are lowered and the other h-links have their weights raised by equal portions of the lowered value (Fig. 6.4). Therefore, the system has now learned that a "square" with a missing edge can still be regarded as a "square".

Regardless of the location of the missing edge, this system would have recognized the object as a "square". It does this by requiring that every cusp activate an h-link with the best possible match to its own parameters and therefore result in the highest total activation value. While this will result in the identical identification of all four missing-

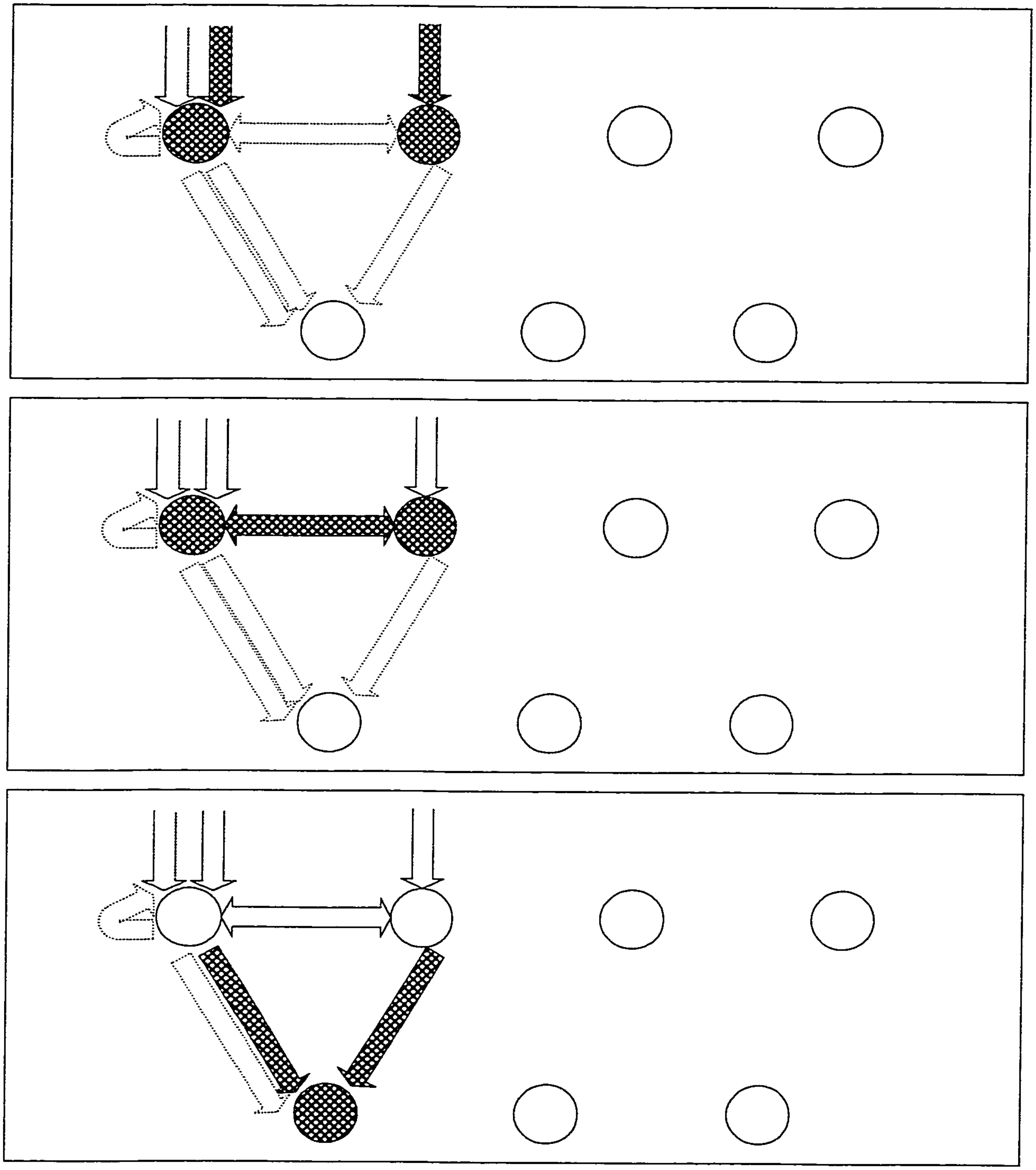

Figure 6.4 Demonstration object three
Partial match of the v-links will cause the pattern of activation to follow a previously learned path.

edge “squares”, it does provide a form of partial rotational invariance that is important when applying the system to more complex objects. The process for searching for the best match requires a longer time than with a previously object, therefore this process can be seen to simulate the increased amount of time for internally rotating the object until it is recognized.

#### 6.2.2.4 OBJECT FOUR

This object is also a right angle square but at a larger size than object one and with a missing half-edge. This missing half-edge is treated as two partial edges assigned to the two involved cusps, therefore the object consists of two cusps with two complete edges, two cusps each with a single edge and partial edge and two cusps each with a partial edge and pseudo edge. Because of the size difference, the “square” v-links are not reactivated for the cusps with complete edges, rather they are assigned to new v-links. Along with the two cusps with single edges and partial edges they are assigned to the same c-cells as the “square” object because of the similarity of the angles between the edges. The two other cusps with only a single partial edge each are assigned to a new c-cell, since they have no angle similarity with previously learned objects.

The intraconnection of the cusps assigned to the two previously used c-cells is similar to that of object three. The two cusps with complete links make use of the previously existing i-links and activate them while adjusting their distance parameters. The other two cusps also in the same c-cell pair reactivate a single i-link each corresponding to their complete edges and generate two new i-links to connect to the cusps in the new c-cell. The cusps in the new c-cell each generate a single i-link for their partial edges and connect to their neighboring cusps in the c-cell pair. In the propagation phase, the previous h-link path is reactivated its source c-cells have been also activated, this results in the four h-links being reactivated with modified distance values to compensate for the larger size of the object and two null distance entries for the two cusps that had partial edges connecting to cusps in a different c-cell. This h-link activation results in a scaled activation of the h-link entry and upon confirmation of the "square" label, the h-links have their weights adjusted as in object three (Fig. 6.5).

The two cusps assigned to the new c-cell did not play any role in the recognition of the object. This is because "square" recognition is satisfied with only the three complete edges of the object, therefore the system has not acquired any new knowledge at the level of classification. However, if the recognition was refuted by the mentor, a new h-link entry would be established and connected to it would be four h-links from the "square" c-cell pair in addition to two h-links from the new c-cell representing the two cusps. This would result in the system learning a new shape, that contains the same components of "square", but also has other additional components.

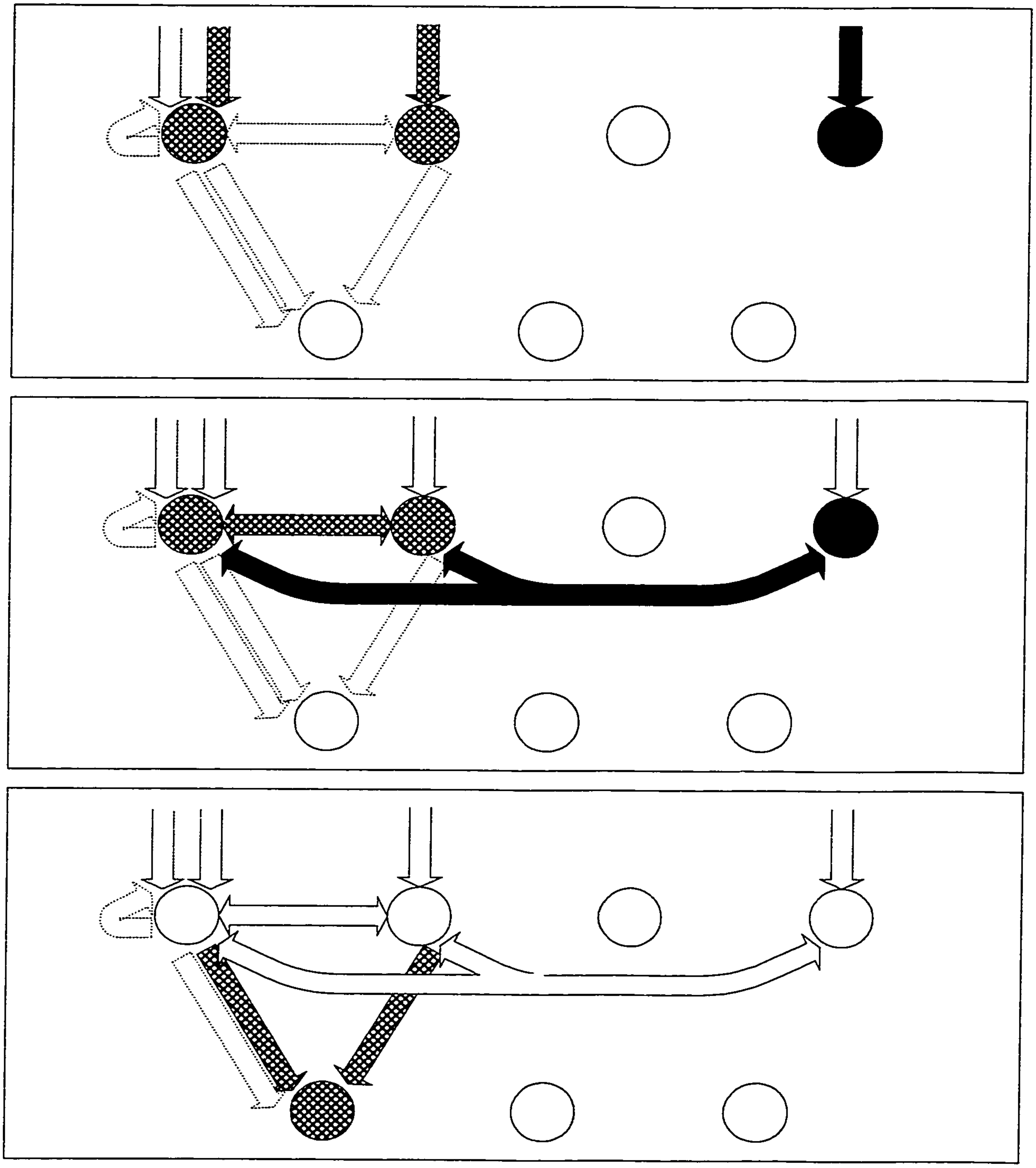

Figure 6.5 Demonstration object four
V-links are assigned to a previously learned path and a new c-cell. Since object classification didn't require input from the new c-cell, no new h-link path was generated.

## 6.2.2.5 OBJECT FIVE

This object has edges of uneven length. Therefore, upon cusp extraction new v-links are generated because the previous v-links while having similar edge directional angles have different size edges. The similarity in edge directions also causes the v-link entries to be assigned to the same "square" object c-cell pair.

The existing i-links between the two c-cells are reactivated with modified edge distances, with no need to generate any new i-links. The h-links corresponding to the activated c-cells are also reactivated, but the destination h-link entry is not activated. This is due to the uneven scaled values carried by the h-links, resulting in the "square" label tolerance being unsatisfied. Therefore, new h-links are generated to a new h-link c-cell entry and the user provides the label "rectangle" (Fig. 6.6).

If desired that this object also be recognized as a "square", upon mentor reassignment the newly generated h-links and their entry would be reassigned to the c-cell holding the "square" label. The logical interpretation of this operation would be that both objects are squares, but object five is viewed from a different perspective than object one. However, it may have occurred that this object satisfied the "square" label tolerance and was subsequently assigned. If this assignment was refuted by the mentor, the e-links would accordingly alter the erroneous assignment depending on whether it was found to be a perfect or scaled match.

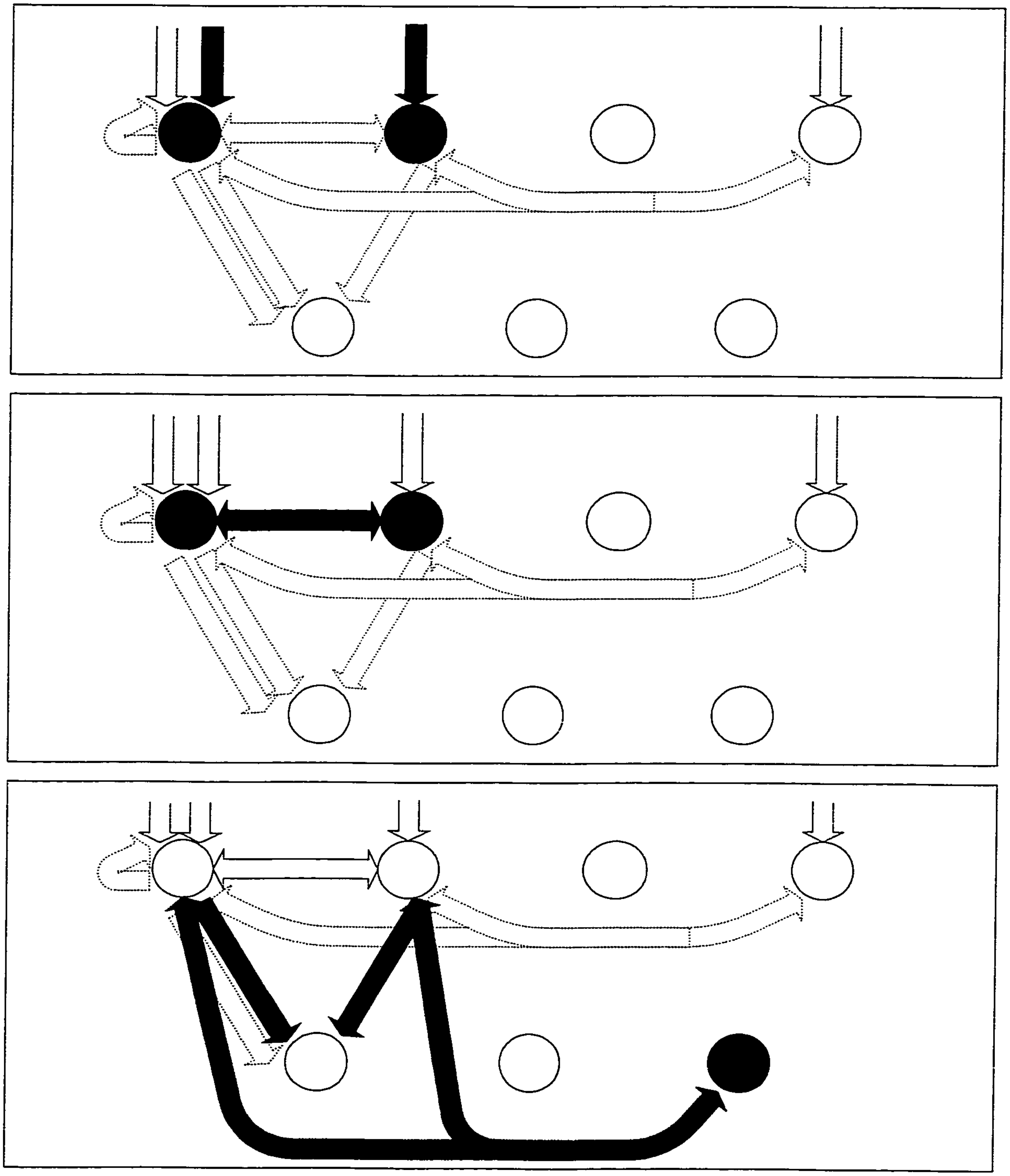

Figure 6.6 Demonstration object five
V-links and i-links follow a previously learned path. Due to incorrect classification, the object has new h-links generated to a new c-cell.

### 6.2.3 ANALYSIS OF EXPERIMENT

These previous experiments demonstrated the basic abilities of the system pertaining to simple object recognition. The same capabilities are used by single objects that are more complex as well as multi-object applications. Limited noise or rotation affecting the input images would not greatly alter the capabilities of the system. However, with the small objects used in this experiment only 10-15 pixels wide, even the smallest amount of noise would be considered a feature, since the filtration window would only be 1-2 pixels wide. For such a system with only a limited number of c-cells, links simply connect to any suitable destination c-cell regardless of its spatial locality to the current c-cell. This connection policy could be later modified as the number of c-cells increases to consider spatial location in choosing a destination c-cell, in the interest of efficient search times, spatial locality and greater biological conformity.

## 6.3 HIGH-LEVEL COMPLEXITY APPLICATION

This experiment will exemplify the recognition of complex objects consisting of multiple components. It will demonstrate the required capabilities of performing such a task through the analysis of two images.

6.3.1 EXPERIMENTAL SETUP

This experiment consists of two multi-component images, with the goal of first recognizing the components of each image then combining them into a single object representing the image. The images to be used are shown in (Fig. 6.7).

As in the previous experiment, each object will be followed by a figure describing its operations (Fig. 6.8,6.9). The first part of the figure represents the operation after the h-links of the simple components have been generated. The middle part illustrates the intraconnection stage involving the simple objects, and the final part illustrates h-link generation of these simple components into a single final object.

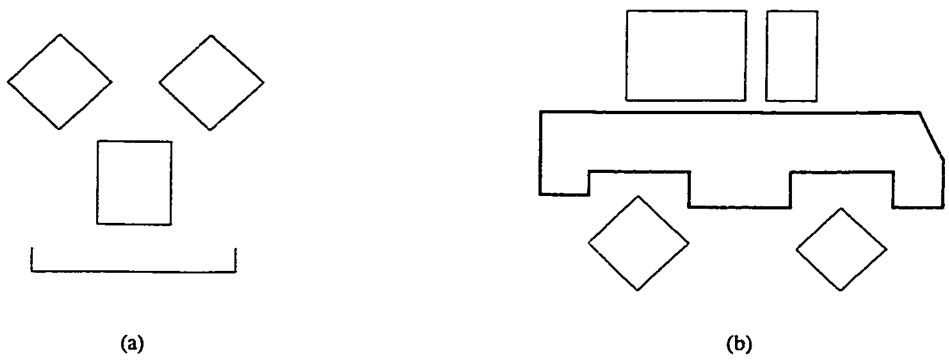


Figure 6.7 Objects involved in high-complexity application
Object (a) was first applied to a new system, followed by object (b). Both objects are enlarged for clarity, but still to scale.

## 6.3.2 EXPERIMENTAL RESULTS

The following experimental results clearly demonstrate the underlying parallel process. Since the system is simulated on a serial machine, this placed significant importance on the blocking mechanisms used to synchronize operation.

### 6.3.2.1 OBJECT ONE

This is a complex object consisting of five simple objects (Fig. 6.7-a). It is a simplified face containing two eyes, a nose, and a mouth. Upon starting with a new system, the entire image is introduced as a single point of focus. The system will then proceed to detect and learn the simple components in a pseudo-parallel manner and combine them into a single object. Operation is exposed in (Fig. 6.8).

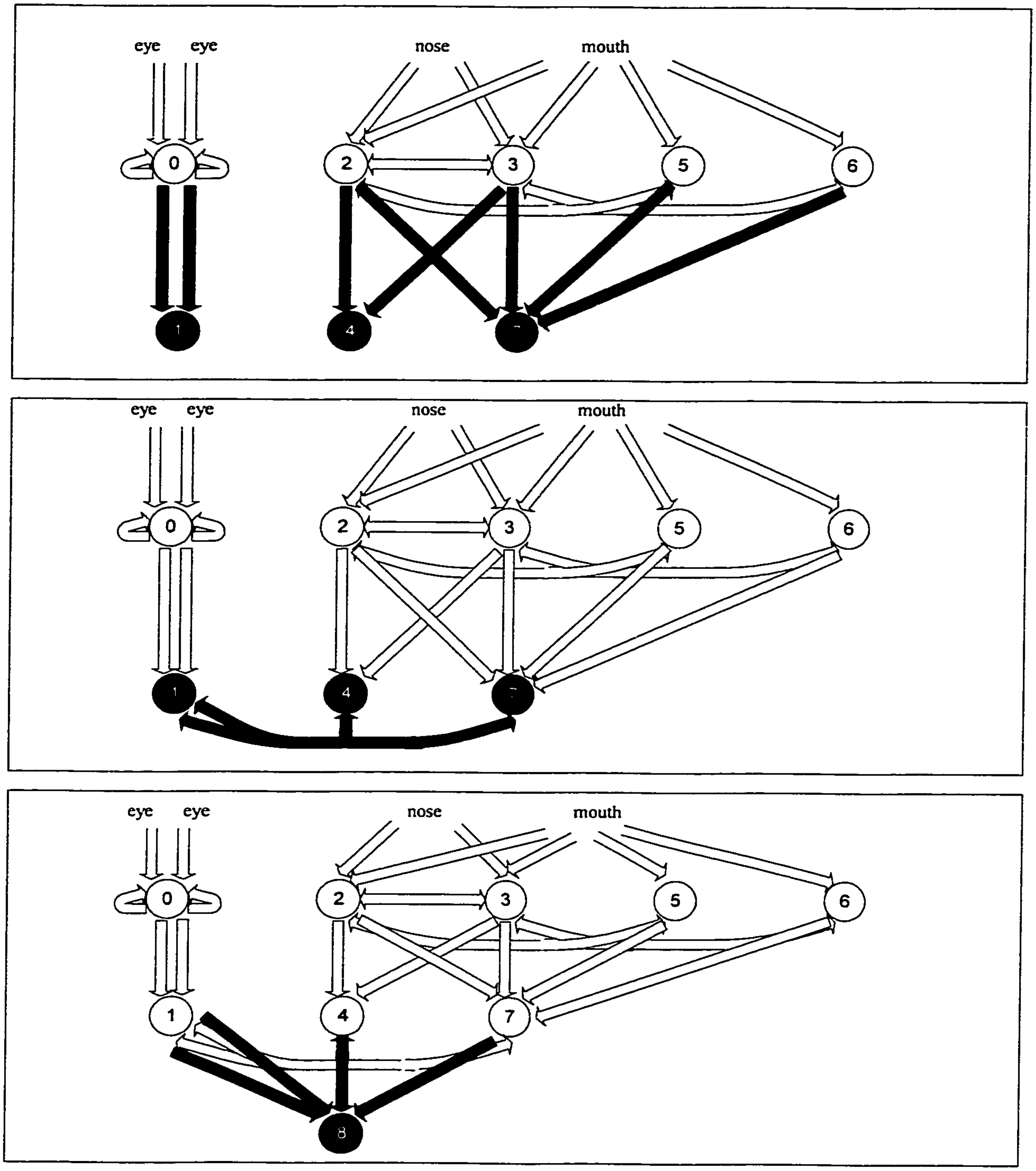


Figure 6.8 Application object one
Upon assignment to a new system the primitive shapes are first analyzed and classified (top), then intraconnected (middle), and finally classified as a single object (bottom).

Since this is a serial implementation of a parallel process, every simple object is detected and learned sequentially before being combined into the complex object. The first object to be processed is the left eye. Learning of this object is similar to that of object two of the low-complexity experiment. After the mentor assigns the label "eye" to this object, the connections are finalized and this object has reached steady-state. Since more objects exist, the label c-cell changes to state-3. This state is not native to the system architecture and is only used in this simulation to perform pseudo-parallel operation. The second object to be detected is also an eye and its v-links are assigned to the same c-cell as of the previous object. Upon i-link generation, existing compatible links are found to exist and are subsequently cloned and have their internal data modified to the new object. A second set of default-mode i-links are also generated between the c-cell v-link entries. Since i-links were cloned, they will follow the same h-link path of their originals. The h-link path is found to already be in use so it is also cloned before being activated. In addition, a default h-link path also propagates from the default i-link entries. The cloned path is found to perfectly activate the destination c-cell h-link entry, resulting in the deletion of the default path. The system announces to the mentor that it has discovered another eye and this discovery is confirmed and the c-cell remains in state-3.

The third object is similar to object five of the low-complexity experiment and its activation path is orthogonal to that of the previous two objects. It is an unknown object assigned the label "nose" by the mentor and consequently goes to state-3. The final simple object is the "mouth", it is similar to object three of the low-complexity experiment in that two of its cusps have only a single edge each. Due to the peculiarities of the low-level image acquisition module, its corresponding filtration processes and most importantly the distortion of small images due to digitization, two cusps will be at right angles while the other two will be at slightly acute angles. This will result in the right angle cusps being assigned to the "nose" object pair c-cells and the other two to new c-cells. Upon i-link generation, the right-angle cusps will find existing compatible i-links and clone them, while the entire object will generate a default i-link path among its four c-cells. The cloned i-link path will generate two cloned h-links corresponding to the number of cusps and the default path will generate four new h-links directed toward a new c-cell. The activation of the cloned path is found to be insufficient for activation of the h-link entry and is therefore deleted. The default path then has its new h-link entry activated and upon mentor query is assigned the label mouth.

At this point all four objects have been recognized and it is time for the second sequential stage. All the c-cells in state-3, are reactivated into state-1 and become ready for intraconnection. These four objects can be thought of as four new cusps that have imaginary edges between them. Since no previous path exists, a default i-link path is generated to fully connect the four objects. In contrast to a simple connected object where only the real edges remain while the imaginary ones are simply not generated, with an unconnected complex object all the edges are kept. Upon the completion of intraconnection, default h-links are generated from the i-link entries and converge into a single new h-link entry. This new entry is assigned the label "face" by the mentor and since all the c-cells are now inactive the system has reached its final steady-state.

### 6.3.2.2 OBJECT TWO

This object is introduced to the system following object one, therefore the system is familiar with the concepts of eye, nose, mouth and face. Object two, represents a simple car with two unequal windows, a rather complex body and two unequal wheels (Fig. 6.7-b). The analysis will follow the same pseudo-parallel technique of the previous object and its operation will be exposed in (Fig. 6.9).

The first component to be analyzed is the left window, it is assigned to same c-cell pair of the nose/mouth objects and upon h-link generation to both mouth and nose labeled c-cells, the object is found to be a scaled version of the stored concept of nose and the corresponding h-link entry is activated. The system announces to the mentor that it has discovered a nose. Since the nose component of object one and the window component of this object are indeed quite similar, the findings are confirmed. If they were refuted the default path would have been activated, thereby assigning this component to a different c-cell. The second window also activates the same c-cell pair of the nose/mouth objects in addition to the default i-link path. Upon h-link generation of the cloned path, the h-link entry for neither nose or mouth are not activated due to the edge lengths of this object that are not equal or even a scaled version of the nose object and an even weaker attempt at activating the mouth h-link entry. Therefore, this cloned path is deleted and the default path converges to a new c-cell. Upon mentor query, it is assigned the label "nose". The system finds that this label already exists and so this object must be an alternative view of the nose class and therefore moves its h-link entry from the new c-cell to the existing "nose" label c-cell. An alternative way of analyzing the classification of this last object was that it was an example of multidimensional classification. Classification in the first dimension occurred with the matching of the angular directions of the edges and was found to match that of both the nose and mouth object classes. However, the second classification dimension dealing with the specific edge lengths found a highly erroneous match to the mouth object and a much closer but nevertheless, erroneous match to the nose object, and thus the result was an incomplete match to both the nose and mouth classes.

The following component to be analyzed is the car's body, it is quite a complex component with five of its v-links being assigned to the "nose" object c-cell pair due to the right angles and the remaining five v-links being assigned to new c-cells. This assignment is primarily due to the tolerance threshold of the system, the higher the tolerance value the more likely the v-links would have been assigned to a fewer number of c-cells and vice-versa. Upon i-link generation, a default path is established between these ten v-link entries, in addition to the v-link entries assigned to the nose/mouth c-cell pair cloning a nose object path and reactivating the dormant mouth object path. H-link generation results in the activation of three paths, the first consisting of two h-links to the mouth object entry, the second is a single h-link to the nose object entry, and the third consisting of the default path connecting all ten i-link entries to a new c-cell. The first two h-link paths are insufficient to activate their corresponding entries and therefore have their entire paths deleted in the case of the cloned nose object path and deactivated in the case of the mouth object path. The remaining default path is activated and announced to the mentor as a new object, which is assigned the label "body".

The last two components to be analyzed are the wheels. The left wheel reactivates the "eye" object path and upon h-link generation is found to give a perfect match to the h-link entry and is subsequently confirmed by the mentor. The second wheel also reactivates the second dormant "eye" object path and although its edge lengths are smaller than those of the original, the h-link entry is activated through a scaled match. These findings are also confirmed by the mentor. Therefore, all the connected components of this object have been analyzed and are in the state-3 synchronization state.

The final car object can now be said to consist of two nose shapes, one body shape and two eye shapes. Their intraconnection involves the generation of a default path connecting all the components and the reactivation of the dormant path intraconnecting the left window "nose" with the two wheel "eyes". The emerging h-links of the reactivated path do not activate the "face" object due to a combination of factors concerning the imperfect nose and eye sizes, missing mouth object and most importantly the spatial distribution of these components. Therefore, this reactivated path is deactivated and the default path is assigned to a new c-cell that is assigned the label of "car" by the mentor.

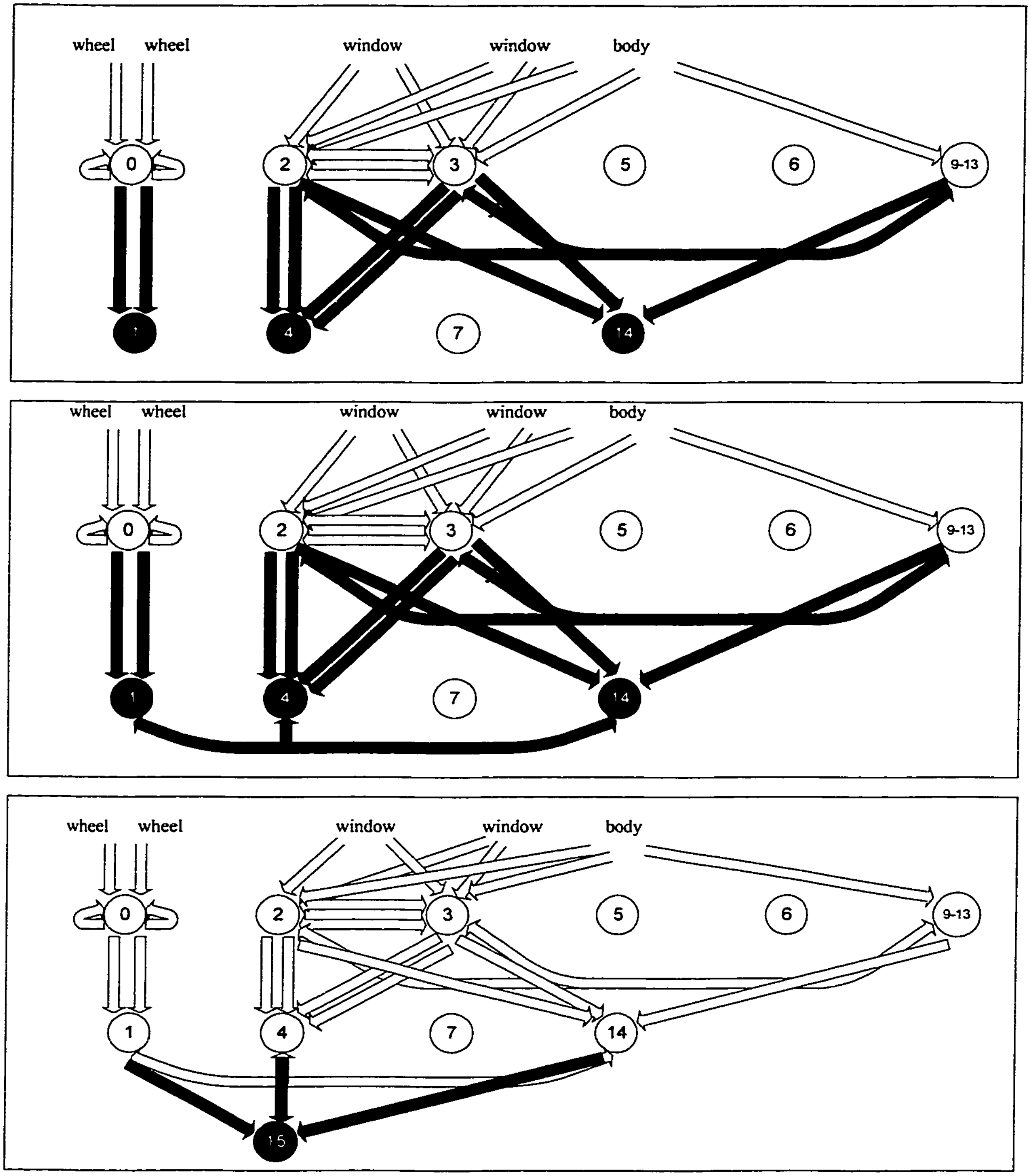


Figure 6.9 Application object two
Assignment of this object's shapes is influenced by the previously learned object. Although some components of the previously learned object exist, they are not sufficient to recall the previous object's classification.

### 6.3.3 ANALYSIS OF EXPERIMENT

This experiment demonstrated the complexity in dealing with objects that consist of more than one component. Also demonstrated was the ability to recognize an object, by recognizing some of its components. For example, if object one was later introduced to the system, but with a single missing component or slightly different spatial placements it would still be recognized as a face, but with a tolerable degree of error. The mentor could then make sure that only a face with a complete set of components is recognized by refuting the result and thereby raising the tolerance of the h-link entry. On the other hand, if the result is accepted, the h-link entries tolerance would not change. This also demonstrates the importance of choosing the default tolerance in calibrating the systems behavior.

## 6.4 CONCLUDING REMARKS

The goal of these experiments was to demonstrate the fundamental capabilities of the system. Similar performance can be expected with objects that are more complex. The system is currently limited to only two levels of hierarchy and a single multi-component object in each image. While, the mechanisms for allowing much higher levels of complexity are supported by the current simulation program, they are not being used. The reason for this is that the system considers each image as single point of focus, therefore all objects in that point of focus are considered part of a single object. In order to support multiple points of focus, a mechanism must be established to allow the differentiation of multiple groups of objects in a single image. As to the apparent limitation of only two levels of hierarchy, if more than one multi-component object was available for analysis, higher levels of hierarchy would be displayed. Possible solutions to this problem and a general extension of the systems capabilities are outlined in the next chapter.

# CHAPTER SEVEN

# EXTENSION OF SYSTEM FOR ADVANCED VISUAL APPLICATIONS

## 7.1 OVERVIEW

As mentioned in the system description, the CSSC machine is designed to be an abstract robust framework with expansion in mind. The previous two chapters described a primitive vision application and its corresponding experimental results. In this chapter, an overview of how the system can be upgraded to the level of advanced vision will be explained.

In order to achieve advanced vision, several additional modules must be added to the present system. These modules will serve the different advanced requirements of such a system, while being based on the same core system principles.

## 7.2 EXTENDED VISUALIZATION MODULE

This module will provide for a direct extension of the visual processing capabilities of the system. Enabling the system to make use of the more complex visual data offered.

### 7.2.1 DESCRIPTION

The current exemplar system is capable of performing basic visual tasks with rudimentary two-dimensional monochrome images. In order to allow for advanced imaging applications, several additional factors must be taken into consideration. With the enhancements offered by the addition of this module, the system will be able to deal with images consisting of several layers of hierarchical objects, employ improved curve fitting to better capture the details of input images, and be able to operate with fully three-dimensional input images.

### 7.2.2 REALIZATION

In order to analyze multiple layers of hierarchical objects, the system must be implemented over a parallel architecture that will allow for efficient analysis of several points of focus simultaneously. The points of focus could be determined by the extremity of the contrast difference between objects in the image, or even by attempting to track dynamic objects. This parallel analysis of the input image is unquestionably performed by the biological eye. An example of this is while reading, the practiced reader will most likely sequentially analyze a document by reading several words in parallel and understanding them before moving on to the next set of words. Therefore, this capability is extremely important in an advanced visual system and would enable much more complex and vivid levels of image description than the two levels currently allowed by the basic system.

The approximation of edges to curves would greatly enhance the system's ability to deal with complex objects. Standard curve-fitting algorithms could be used in capturing the objects contour.

The final enhancement would be to allow for full three-dimensional imaging. As mentioned, two-dimensional projections are sufficient for such imaging. However, several other factors must be taken into consideration to fully enable this feature. First, in order to differentiate between an image that is two-dimensional and one that is a projection of a three-dimensional image, visual cues must be considered. These visual cues can frequently be seen in static animations, where an attempt is made to "trick" the eye into believing that it is viewing the image at a certain perspective (Fig. 7.1). If the system was intended to work with dynamic images in a three-dimensional setting then the camera point would be the visual cue as to the perspective of the image viewed [33], [34]. The importance of being able to process several points of focus is again exemplified for this enhancement.

These modifications would for the most part add only additional data fields to the connection links, their processing would remain largely the same except for the availability of more detailed information. In conducting several operations of analysis in parallel, only a single copy of the cells and links would be required that would block/grant read/write access according to the circumstances of the simulation.

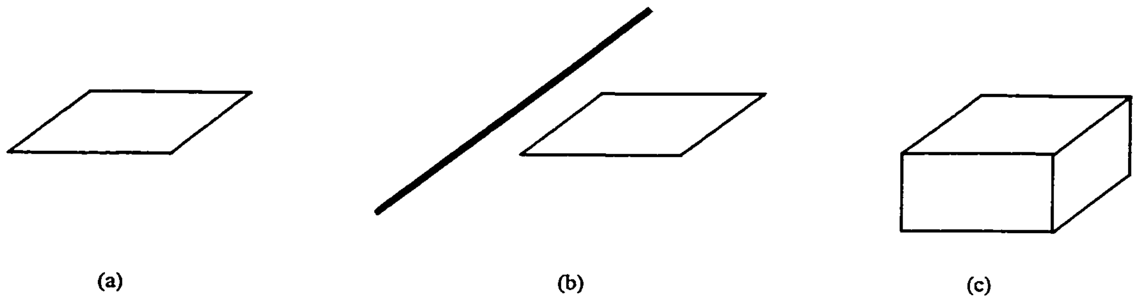


Figure 6.1 Visual deception
(a) Object is a normal parallelogram
(b) Same object next to high contrst line gives the illusion of depth
(c) Same object connected to other objects gives the illusion of depth

## 7.3 LINGUISTIC PROCESSING MODULE

This module's main purpose is to homogeneously communicate with the mentor and translate the top-down instructions into a form consistent with the systems native data formats.

### 7.3.1 DESCRIPTION

Functionally, the differences between vision and language are profound [35]. The role of vision is to keep us informed about the world we inhabit, whereas language is primarily used to exchange experience and collaborate with others. However, it is now largely accepted that there are many analogies between visual and natural language processing, whether they contain some common features, or that they completely share the same semantics, is still open to debate. However, the following characteristics are common between them:

Hierarchical organization.

In language, it goes down from entire sentences to words to letters. Similarly, in vision we start from a single image to objects to simple shapes. Interaction between levels of the hierarchy is a function of the degree of restriction of context. In a highly restricted sense, a change in a single letter can change an entire language script; likewise, a change in a single shape can redefine an object.

Syntax-semantics division.

For language, syntax refers to the way in which words are put together to form phrases and sentences, while semantics pertain to the actual meanings of the resultant phrases and sentences [5]. Vision has a similar division, where objects that are correctly constructed from known primitive components do not always result in the final objects being identified.

Ambiguity

In language, it can be referred to as misunderstanding due to the use of unusual forms or grammatical cues. Visually such ambiguity is commonly called an optical illusion due to an unusual viewing perspective or the use of salient visual cues.

The previous list of common characteristics among vision and language may well indicate a deep relationship them, which could be used in designing a system that would uniformly process both types of data. This could also mean that there is an analogous syntax and semantics of vision. Ultimately, this could very well mean that language and visual processing are performed in an analogous way inside the brain with only minor alterations to a common abstract model. We could also consider knowledge as physical experience (syntax) that is processed into conceptual operational knowledge (semantics), and thereby potentially have a unifying theory of vision, language, and knowledge.

Adding a linguistic processor to the system would at very least provide a homogeneous method of interacting with the outside world. If fully implemented, the result would be the ability to translate top-down instructions from the mentor into a homogenous consistent form usable by the system.

### 7.3.2 REALIZATION

Implementation of such a module could follow the same procedure as that taken when developing the current system. These procedures would consist of the aggregation of information into the actual workings of the actual process in humans and finally, an attempt at its simplification, generalization and ultimately simulation. However, the design would not necessarily have to begin from scratch, since the current system framework was designed to be of sufficient robustness to allow for new applications. In addition, the ultimate strength of the relationship between visual and linguistic processing could reduce the development effort by reusing many of the visual concepts already developed.

## 7.4 LOGIC PROCESSING MODULE

This module is designed to independently construct complex relations from stored data, effectively it would operate as a trainable functional unit. While it may first seem incorrect to separate complex logical processing abilities from the basic capabilities of the underlying CSSC machine components. There is a growing body of evidence that biological synapses can be adjusted directly by extrasynaptic neuromodulators, thereby removing the restriction of local learning laws as the only means of connection adjustment [29]. This would allow for empirical local adjustment of connections and rule-based (mentor) adjustment of connections by other external modules of the CSSC machine.

### 7.4.1 DESCRIPTION

This is the least understood and most complex module. If successfully developed, the resulting system would be able to completely process input data and learn from examples using the vision processing model, interact and learn from its mentor using the linguistic processing module and be able to self-construct complex logical thought using this module.

The true ramifications of the development of such a module are immense and could potentially be used by any other application developed for this system. This module would enable the aggregation of knowledge and combining it with discovered or taught theories. The result would be the ability to achieve goal-oriented searching and the construction of complex relations between the data as required by high-level symbolic processing. Effectively by understanding concepts and establishing relations between them, the system would be able to "think" [36], potentially produce novel solutions and generate basic forms of "creativity" [37].

### 7.4.2 REALIZATION

Despite the difficulty of construction of such a module, the basic framework developed thus far should help in its realization. Since this framework logically stores data in a structured manner, this would simplify the task of correct data extraction. A new class of links can be developed that would interface between the linguistic module and other perceptual modules. This new link class would be able to directly adjust individual link entries, thereby allowing for rule-based learning. In addition, links could be designed to internally adjust other links and entries to allow the construction of more complex relations among stored data.

It would be a mistake to simply reuse the logical processing modules found in symbolic AI systems, since they are not correctly based on actual biological systems and would result in a disparity between the processes and data used by the modules. In addition, such primitive modules are still not functionally sufficient to consider their use.

## 7.5 CONCLUDING REMARKS

The enhancements discussed in this chapter can all be based on the same basic CSSC machine. The first module would be the least demanding to implement, since its requirements are largely concerned with the front-end image acquisition component of the system. The linguistic processing component would have to be built in the same manner as that followed in the construction of the current basic vision system. While language from the onset is more complex than basic vision, many of the same design elements can simply be reused with only minor modification. Again, this is subject to how close language and vision processing are actually found to be. The most difficult module to construct would be the logic processor. However, it is potentially the most beneficial, since it would be able to independently manipulate stored data and enable the construction of complex relations.

The single most important enhancement to the system's performance would be the availability of an associative memory. Associative searching is a primary operation in the program simulation and the availability of such a memory would increase the speed of the operations tremendously. In addition, it would allow for inherent associability between components of an image, thereby reducing the data storage requirements even further.

## 7.6 FINAL SUMMARY

The result of this work is the introduction of a novel AI architecture. Its novelty is primarily a consequence of its inception on the complete scope of natural intelligence. This design approach allows it to encompass the features of both symbolic and connectionist architectures, while allowing for new medial abilities that are not supported by current AI architectures or their hybrids.

The current abstract system, while exemplified for basic vision, can support other perceptual applications. However, when implementing this system for any perceptual process, data supplied to the system and any additional types of links must be carefully chosen to comply with the natural perceptual process in order not to nullify the benefits of the architecture. The ultimate potential of this architecture will only become apparent after the construction and integration of several perceptual modules, after which the original design principles can be fairly evaluated for their suitability.

## REFERENCES


[1] Brooks, R.A. "Intelligence Without Reason", MIT A.I. Memo No. 1293, April 1991.

[2] Rumelhart, D.E. and McClelland, J.L. ed. Parallel Distributed Processing Vol. 1. Cambridge (MA): MIT Press, 1987.

[3] Honavar, V. and Uhr, L. ed. Artificial Intelligence and Neural Networks. San Diego (CA): Academic Press Inc., 1994.

[4] Way, E.C. "Conceptual Graphs Past, Present, and Future." Second Int. Conf. on Conceptual Structures, p.11-30, Baltimore (MD), 1994.

[5] American Heritage Dictionary. 2nd ed. 1991.

[6] Levine, M.D. Vision in man and machine. N.p.: McGraw-Hill Book Co., 1985.

[7] Dayhoff, J.E. Neural network architectures. New York (NY): Van Nostrand Reinhold, 1990.

[8] Rumelhart, D.E. and McClelland, J.L. ed. Parallel Distributed Processing Vol. 2. Cambridge (MA): MIT Press, 1987.

[9] Strange, P.G. "Basic Brain Mechanisms: A Biologist's View of Neural Networks." IEE Colloquium on Current Neural Network Research. 1 (1989): 1-5.

[10] Jin, Y. and Zhang, J. "Three-Level Neural Model: An Operational Study of Essential Human Intelligence." Int. Joint Conf. on Neural Networks, 1, p. 747-753, 1989.

[11] Harris, C.S. ed. Visual Coding and Adaptability. New Jersey (NJ): Laurence Erlbaum Assoc., 1980.

[12] Hayes, P. "Aristotelian and Platonic Views of Knowledge and Representation." Second Int. Conf. on Conceptual Structures, p.1-10, Baltimore (MD), 1994.

[13] Poggio, T. and Shelton C.R. "Machine Learning, Machine Vision, and the Brain." AI Magazine. (Fall 1999): 37-55.

[14] Rosenfeld, A. ed. Human and Machine Vision II. Orlando (FA): Academic Press Inc., 1985.

[15] Gregory, R.L. Eye and Brain, the Psychology of Seeing. N.p.: World University Library, 1973.

[16] Feldman, Y.A. and Bruckstein, A. ed. Artificial intelligence and computer vision. Amsterdam: Elsevier science Publishers B.V., 1991

[17] Leavers, V.F. "Shape Detection in Computer Vision Using the Hough Transform", Springer-Verlag, 1992

[18] Mohan, A. "Object Detection in Images by Components." MIT A.I. Memo No. 1664, June 1999.

[19] Winston, P.H. ed. The Psychology of Computer Vision. McGraw-Hill Book Co., 1975.

[20] Wooldridge, D.E. Sensory Processing in the Brain. New York (NY): John Wiley & Sons, 1979.

[21] Ullman, S. and Basri, R. "Recognition by Linear Combinations of Models." IEEE Trans. on Pattern Analysis and Machine Intelligence. 13 (Oct. 1991): 992-1006.

[22] Winston, P.H. Artificial Intelligence. Reading (MA): Addison-Wesley Pub. Co., 1992.

[23] Aamodt, A. "A Knowledge Representation System for Integration of General and Case-Specific Knowledge." Proc. of Sixth Int. Conf. On Tools with Artificial Intelligence, p. 836-839, 1994.

[24] Fu, L. "Knowledge Discovery Based on Neural Networks." Communications of the ACM. 42 (Nov. 1999): 47-50.

[25] Sowa, J.F. Conceptual Structures. N.p.: Addison-Wesley Publishing Co., 1984.

[26] Shavlik, J. "A Framework for Combining Symbolic and Neural Learning", Computer Sciences Dept. Univ. of Wisconsin-Madison Technical Report 1123, Nov. 1992.

[27] Schwinn, J. "A Hybrid Approach for Knowledge Representation and Reasoning." First New Zealand Int. Conf. On Artificial Neural Networks and Expert Systems, p. 214-215, 1993.

[28] Towell, G.G. and Shavlik, J.W. "Knowledge-Based Artificial Neural Networks." Artificial Intelligence. Vol. 69, 1994.

[29] Roy, A. "On Connectionism, Rule Extraction, and Brain-Like Learning." IEEE Trans. on Fuzzy Systems. 8 (Apr. 2000): 222-227.

[30] Gonzalez, R. and Wintz, P. Digital Image Processing. Addison-Wesley, 2nd Ed. 1987.

[31] Freeman, W.J. "A Neurobiological Interpretation of Semiotics: Meaning vs. Representation." IEEE Int. Conf. on Computational Cybernetics and Simulation, p. 1481-1486, 1997.

[32] Wilson, G.R. and Batchelor, B.G. "Algorithm for Forming Relationships between Objects in a Scene", IEE Proc. E. 137 (1990): 151-153

[33] Harrington, S. Computer Graphics a Programming Approach. McGraw-Hill Book Co., 1985.

[34] Jain, R. et. al. Machine Vision. N.p.: McGraw-Hill Inc., 1995

[35] McKevitt, P. ed. Integration of Natural Language and Vision Processing. Dordrecht: Kluwer Academic Publishers, 1995.

[36] Doval, D. "From Neurons to Consciousness." IEEE Int. Joint Symp. on Intelligence and Systems, p. 38-43, 1996.

[37] Boden, M.A. "Creativity and Artificial Intelligence." Artificial Intelligence. Vol. 103, 1998, p. 347-356.

# VITA

Ahmed A. Fadol

Bachelor of Science in Computer Engineering, 1998

King Fahd University of Petroleum & Minerals

Master of Science in Computer Engineering, 2001

King Fahd University of Petroleum & Minerals