Department of Electrical Engineering

Cullen College of Engineering

University of Houston

(NASA-CR-189537) SELF-ORGANIZATION VIA ACTIVE EXPLORATION IN ROBOTIC APPLICATIONS Final Report (Houston Univ.) 59 p CSCL 13I

N92-19289

Unclas

G3/37 0075762

Self-organization via Active Exploration in Robotic Applications

FINAL PROGRESS REPORT NAG 9-473

submitted to NASA-JSC IN-37-CR

by

H. Öğmen, Principal Investigator 73762

R. V. Prakash, Research Assistant p. 59

Department of Electrical Engineering

University of Houston

Houston, TX 77204-4793

February, 1992

Contents

1 Introduction **2**

2 General principles **3**

3 Reinforcement learning and habits for reinforcement filtering **4**

3.1 Neurophysiological basis 4

3.2 Significance 8

3.3 Neural models 8

3.4 Modifications 13

3.5 Simulation Results 17

3.5.1 ART network 17

3.5.2 Reinforcement-habit network 17

4 Novelty and Reinforcement **21**

4.1 Neurophysiological basis 21

4.2 Significance 24

4.3 Neural models 24

4.4 Modifications 28

4.5 Simulation Results 34

4.5.1 Simulations on Gated Dipole 34

4.5.2 Simulations on the new modified model 34

5 The combined model **38**

5.1 Ambiguity in decision making 46

5.2 Combined neural model 46

5.3 Simulation Results 52

6 Conclusion **54**

1 Introduction

Various space applications require the design of autonomous systems capable of perceiving, reasoning, and acting in a complex non-stationary environment. Moreover, in order to achieve a maximum degree of autonomy, these systems should be self-calibrating, i.e. highly adaptive to changes in their structure. In traditional robotics perception, reasoning, and motor control have been treated as mainly independent modules due to the strong reductionism of mainstream artificial intelligence. While various artificial intelligence algorithms have been proposed for these modules, these algorithms have been proved to be inadequate for *non-stationary* environments. For example, vision has been traditionally separated into two parts, preattentive and postattentive vision. The former has been formalized as an "inverse optics" problem while the latter as a symbol manipulation problem. Inverse optics attempts to invert geometric and radiometric equations relating the properties of the three-dimensional environment to the shape and luminance of the image. The inverse problem is ill posed, and traditional computer vision uses additional assumptions to reduce the solution space. However, these assumptions imply strong constraints on the environment. When these constraints are not satisfied the performance of the algorithm deteriorates severly. Therefore, these approaches are inadequate for non-stationary environments. A powerful technique for abstract reasoning is the expert system paradigm. Again, the success of expert systems is limited to stationary environments. In motor control, traditional approaches use a reference model for the plant to be controlled (e.g. arm). Even in the presence of such a model, the inverse problems are ill posed and require additional assumptions. Furthermore, the parameters of the actuator change through time (for example due to fatigue) and traditional robotic systems require human intervention for calibration. In order to avoid such interventions, a non-stationary approach becomes desirable.

A recent approach in robotics attempted to avoid these difficulties by simplifying the desired behavior (instead of the environment as in the approaches mentioned above) (Brooks, 1986, 1989). This enabled the design and hardware implementation of autonomous systems with simple behaviors such as escape and wander. The simplification of the behavior enables one to cross the perception-reasoning-action stages without designing complex structures for each. In fact, the system proposed by these authors is a "hard-wired production-rule system" with simple rules (here simple rules are sufficient because simple behaviors are sought). However, this approach potentially faces the same problems that traditional artificial intelligence faced when it tried to generalize the "block world" techniques to complicated environments, for it does not address the fundamental shortcoming of these approaches: the lack of self-organization tailored for non-stationary environments. In that respect, various neural network approaches that have been proposed for robotics are also inadequate because, while they are capable of self-organization, they cannot cope with non-stationary environments.

In this report, we describe a neural network robotic system designed to self-adapt in non-stationary environments. As we will discuss, active exploration plays an essential role in our approach. This stems from our conceptualization of intelligence as a *process* rather than a fact. Consequently, the understanding and the emulation of intelligent behavior requires the characterization of its dynamics rather than just its equilibria. This led us to study minimal systems with active exploratory capabilities. The system has been implemented and tested in software. We also describe the details of various implementations.

2 General principles

Unlike many primitive animals which are almost completely genetically wired, human infants undergo an extensive developmental period, during which they learn to control and coordinate various parts of their body. This self-organization process is highly adaptive for it is able to control a non-stationary system (e.g. the growing child's arms become longer etc.). Demonstrations of these effects in adults are quite old. Helmholtz showed that adults can adapt to inverting prisms placed in front of their eyes. While many interpretations of this adaptation process have been based on an adult error detection and correction behavior, Held and co-workers defended the view that a single mechanism is responsible for both infant sensory-motor coordination and adult sensory-motor adaptation(Held and Bossom, 1961; Held and Hein, 1963; Held, 1965). Furthermore, their experiments suggested that an *active* control of muscles is necessary for sensory-motor adaptation. The role of active processes as a basis of self-organization goes beyond sensory-motor coordination. Piaget's studies suggested that *active exploration* plays an essential role in the development of intelligence in addition to perception and motor control(Piaget, 1963, 1967, 1969, 1970). Within this framework these three aspects of behavior develop together by continuously being influenced by and influencing each other.

Given this primary role for active exploration in intelligent behavior, the fundamental question that we posed is "what are the simple circuits and organizational principles that are necessary for the initiation of an exploratory behavior ?" To answer this question, we analyzed the first step of exploration, reaching out for targets. The system that we designed has sensorial inputs and motor outputs and a "cognitive unit" that coordinates these two. An important step in intelligent behavior is stimulus generalization. To achieve this behavior we introduced categorization circuits. While there are various categorization circuits proposed in the neural network literature, in order to satisfy our requirement of non-stationary environments, we chose adaptive resonance theory architectures that are capable of forming stable categories in nonstationary environments(Carpenter and Grossberg, 1987, 1988). Simple adaptive resonance circuits require that the inputs are pre-processed by a figure-ground segregation

network. To achieve a very simple figure-ground segregation, we introduced a non-homogeneous retina consisting of a high resolution fovea and a periphery. The figure-ground segregation is achieved by directing the fovea to parts of the image such that the foveal signal is treated as figure and the peripheral signal as ground. This in turn required that we introduce a circuit for controlling eye movements. The basic circuit for eye movements should be flexible to operate under both sensorial and cognitive control to allow an exploratory behavior that accommodates both sensorial and cognitive cues. For that task, we modified some neural circuits proposed for optokinetik behavior in lower animals as the seed of the eye movement system(Öğmen and Gangé 1990a, 1990b). An important characteristic of this circuit is its sensitivity to novelty which is also crucial for exploration. In addition to sensitivity to spatial novelty, the system should be able to recognize the novelty of abstract stimuli representations (categories). Another important property for exploration is the success of behaviors. In simple developing systems the success of behaviors are determined by reinforcement signals and thus the system should be capable of operant conditioning. However like any signal in an uncontrolled environment, reinforcement signals can also be noisy. In order to filter the transients in the reinforcement signals another property is required: habits. In sum, a simple neural network should have stimulus generalization, novelty, reinforcement learning and habit formation properties. Whether these properties are sufficient is the question we attempt to answer by building and analyzing neural network architectures having these properties. The complete systems contains various interacting sub-architectures. We will first start describing these architectures and the modifications to each architecture for the needs of the present application. We will then present the combined model that integrates these architectures into a global system. We finally conclude by discussing future directions in this research.

3 Reinforcement learning and habits for reinforcement filtering

3.1 Neurophysiological basis

Assimilation and avoidance of a behavior is highly correlated with reward and punishment signals. Animals and humans generally tend to favor rewarding events to punishing ones. Moreover, in certain cases the absence of reward could act as punishment and vice versa i.e., not receiving punishment can act as a reward. Studies by Milner and Pribram have suggested that the frontal lobes in primates and humans (see Figure 1) are primarily responsible for correlating the reward and punishment to actions(Milner, 1963, 1964; Pribram, 1961). Milner found that patients with frontal lobe lesion demonstrated difficulty in shifting their response to changing environment. For example if a certain action (like eating sweets) which was formerly rewarding but later turned to be punishing (caused a stomach ache due to overeat-

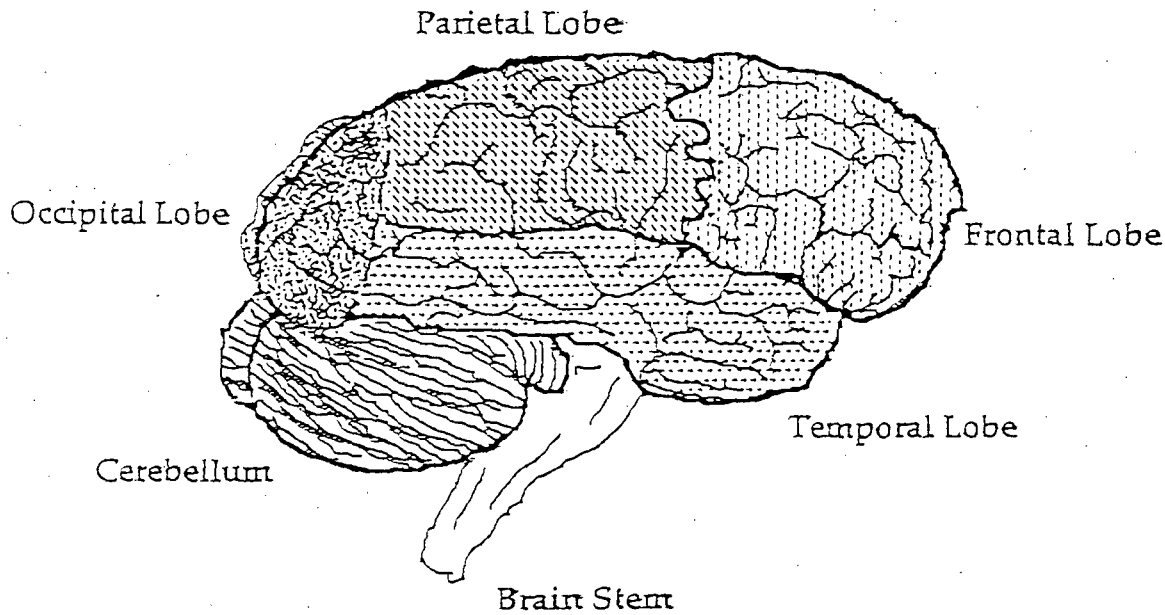


Figure 1: The human brain can be divided externally in to four lobes, the *occipital lobe* (responsible for visual activity), the *parietal lobe* (responsible for abstract thinking e.g logic, math *etc.*,), the *temporal lobe* (responsible for hearing and memory) and the *frontal lobe* (responsible for goal directed behavior, articulation of speech and control of discrete movement of the body).

Locus of lesion	No. of cases	Categories achieved	Total errors	Per-severative	Nonper-severative
Dorsolateral frontal	25	1.372	75.44	56.932	18.5
Control: Orbitofrontal + temporal	7	4.9	27.6	12.0	15.6
Inferior frontal	1	6	13	4	9
Bilateral hippocampal	1	4	48	11	37
Posterior cortex	60	4.52	32.96	16.64	16.32

Table 1: Milner's summarized results

ing), then these patients found it hard to change their original behavior. Her test group consisted of 71 patients who were tested both before and after frontal lobectomy (removal of the frontal lobe for relief from some kinds of seizures). Her test group also included 23 patients who had previously received frontal lobectomies (in some cases years ago). The test which she used on these subjects dealt with matching of similar cards and is called the "Wisconsin Card Test". The subjects were shown four "stimulus cards" (see Figure 2) which consisted of cards differing in color, form and number: one red triangle, two green stars, three yellow crosses and four blue circles. The subjects were then given a pack of 128 "response cards" which were a combination of the above criteria, i.e the four colors, forms and numbers. Figure 2 shows one such response card (two red crosses) which corresponds to the stimulus card 1 in color, to the stimulus card 2 in number, and to the stimulus card 3 in form. The subjects were then told to place the response card in front of one of the four stimulus cards, wherever they thought it should go. They were then informed whether their choice was "right" or "wrong". The subjects used this information and tried to get as many cards right as they could. No other cues were given. The subjects were required to sort first by a certain color, all other responses being called wrong. Then once they achieved 10 consecutive correct responses to color, the sorting criterion was changed to form, without previous warning, and color responses were deemed wrong. After the subjects gave 10 consecutive correct responses to form, the criterion was shifted to number and then back to color again. This procedure continued until the subjects had successfully completed six sorting categories (color, form, number, color, form, number) or until all 128 cards had been placed.

Milner found that the subjects who did not have dorso-lateral frontal lobectomy could change their responses as the sorting criterion changed. However subjects who had dorso-lateral frontal lobectomy could sort the cards correctly only for the first criterion and continued to sort according to this criterion even after the criterion was shifted to form by the experimenter. Milner categorized the "wrong"

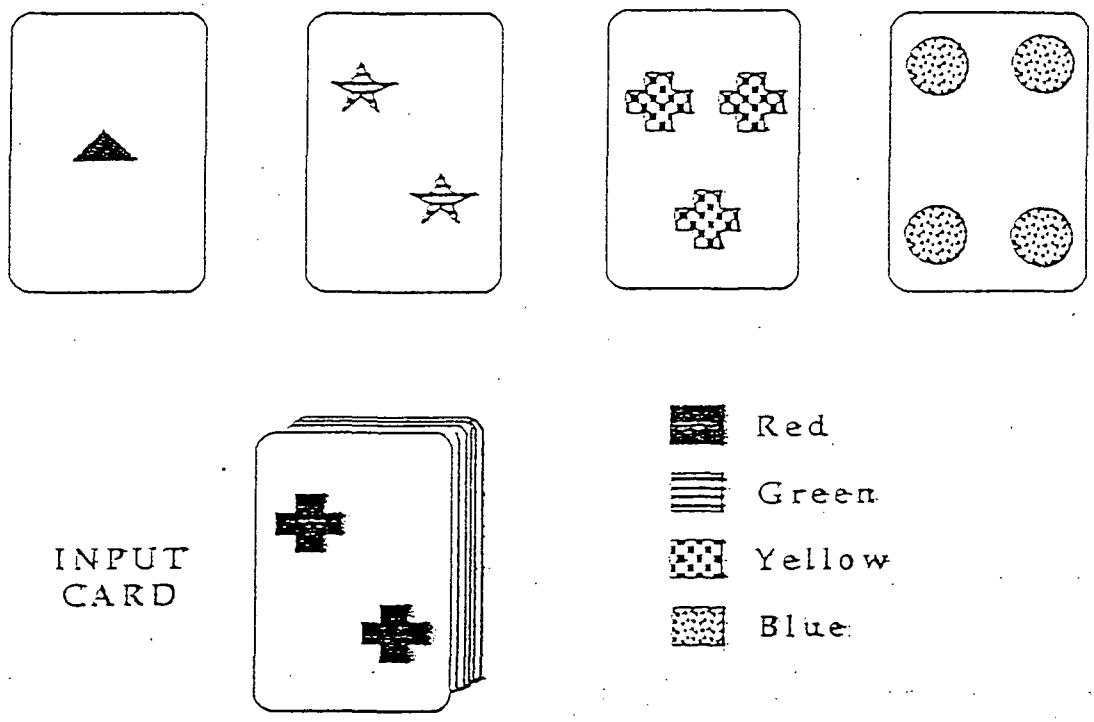


Figure 2: The cards used in the Wisconsin card sorting test. The four cards represent the four "stimulus (template) cards". A sample response card is also shown in the figure

responses given by subjects into two categories. The first was the "perseverative errors" which consisted of responses which would have been correct on the immediately preceding stage of the test, or, in the first stage as a continued response in terms of the subject's initial preference (which is incorrect now). All other errors were lumped together and called as "non-perseverative errors". Table 1 summarizes the findings of Milner. As can be seen from Table 1, subjects with dorso-lateral frontal lobectomy were unable to shift from one sorting principle to another. They committed an average of 57 perseverative errors. Milner reported that in some cases these subjects were unable to shift their criterion throughout the entire experiment. The control subjects on the other hand could change their criterion and had about 15 perseverative errors. Non-perseverative errors in general were not as pronounced as perseverative errors except in the case of subjects who had bilateral hippocampal lesions. This anomaly is attributed to a profound memory loss.

3.2 Significance

Thus Milner's experiment suggests that frontal lobes play a major role in enabling one to shift his/her response to meet changing environment. This is essential for successful performance in nonstationary environment. Yet, as mentioned earlier, reinforcement signals can be noisy and normal humans filter those signals by mixing reinforcement and habit tendencies. Milner's result suggest that if the gain of the reinforcement signal are diminished (due to frontal lobe damage) then habit filters become dominant and cause extreme persevaration of behavior. While Milner's study is directed to anamolous brain function, it reveals an important aspect of reinforcement filtering (habits) which is otherwise transparent to the observer.

3.3 Neural models

Leven and Levine proposed a neural network to model some roles of frontal lobes in the decision making process(Leven and Levine, 1987). Their model comprised of two parts as shown in Figure 3. The left hand side comprises of the categorization network which is responsible for choosing the "correct" category for the given input "response card" and the network on the right hand side comprises of the habit and bias network.

The categorization network comprises of an Adaptive Resonance Theory (ART) network. ART has of two layers: the input layer consisting of feature neurons and the categorization (output layer) consisting of category neurons. Each of the feature neurons codes a particular criterion: color, shape or number of the input "response card". There are a total of twelve neurons in this layer and each criterion is represented by four neurons (one for each of the four attributes in a given criterion). Hence

a particular response card, say two red crosses, is represented by the input vector (0100 1000 0010) where the first four from the left digits represent the number criterion: one, two, three and four, the next four digits represent the color criterion: red, green, yellow and blue and the final four digits represent the form criterion: triangle, stars, crosses and circles respectively. Any response card would activate one of the four neurons of a given criterion. The above input would activate the x[2], x[5] and the x[11] neurons of the input layer.

The categorization layer comprises of four neurons which represent the four "stimulus cards". They are one red triangle, two green stars, three yellow crosses and four blue circles. A given input "response card" would match one of the four "stimulus card" category neurons depending on reward input. Initially, when a "response card" is shown to the categorization network the feature neurons which code the attributes of this response card become active. The activity of a feature neuron is given by the following differential equation.

$$\frac{dx_i}{dt} = -Ax_i + (B - Cx_i)(I_i + \sum_{j=1}^{4} f(y_j)z_{j,i}) - Dx_i \sum_{j=1}^{4} f(y_j), \qquad i = 1, 2, ...12, \tag{1}$$

where I_i is the input vector given to the feature layer (e.g: two red crosses would be represented as $(010010000010) = (I_1, I_2, ..., I_{12}) = \mathrm{I}$), x_i represents the activity of the $i^{\underline{th}}$ feature neuron,y_j is the activity of the $j^{\underline{th}}$ category neuron, and $z_{j,i}$ represents the top down weights connecting the $j^{\underline{th}}$ category neuron with the $i^{\underline{th}}$ feature neuron.A, B, C and D are positive constants and "f" is a sigmoid function which is defined as follows.

$$f(x) = arctan(x - 1) + \frac{\pi}{2} \tag{2}$$

Equation (1) is a shunting equation(Grossberg, 1988) and the activity of feature neurons is bounded (in this case between zero and $\frac{B}{C}$). The first term $-Ax_i$ in the above differential equation is responsible for the passive decay of activity at rate A. The second term $(B - Cx_i)(I_i + \sum_{j=1}^{4} f(y_j)z_{j,i})$ is the excitatory term and it consists of the input I_i and the excitation from the category nodes[1] weighted by the top down weight $z_{j,i}$. The third and final term $-Dx_i \sum_{j=1}^{4} f(y_j)$ is the inhibitory part which comprises of inhibition from the categorization neurons. This inhibition allows the network to distinguish between bottom up and top down signals.

The activity of the category neurons represent the possibility that the input "response card" belongs to that "stimulus card" category. This activity is represented by equation (3).

$$\frac{dy_j}{dt} = -Ay_j + (B - Cy_j)(f(y_j) + \sum_{i=1}^{12} g(\Omega_{[\frac{i+3}{4}]}x_i)w_{i,j})$$

[1] "node" and "neuron" have been used synonymously

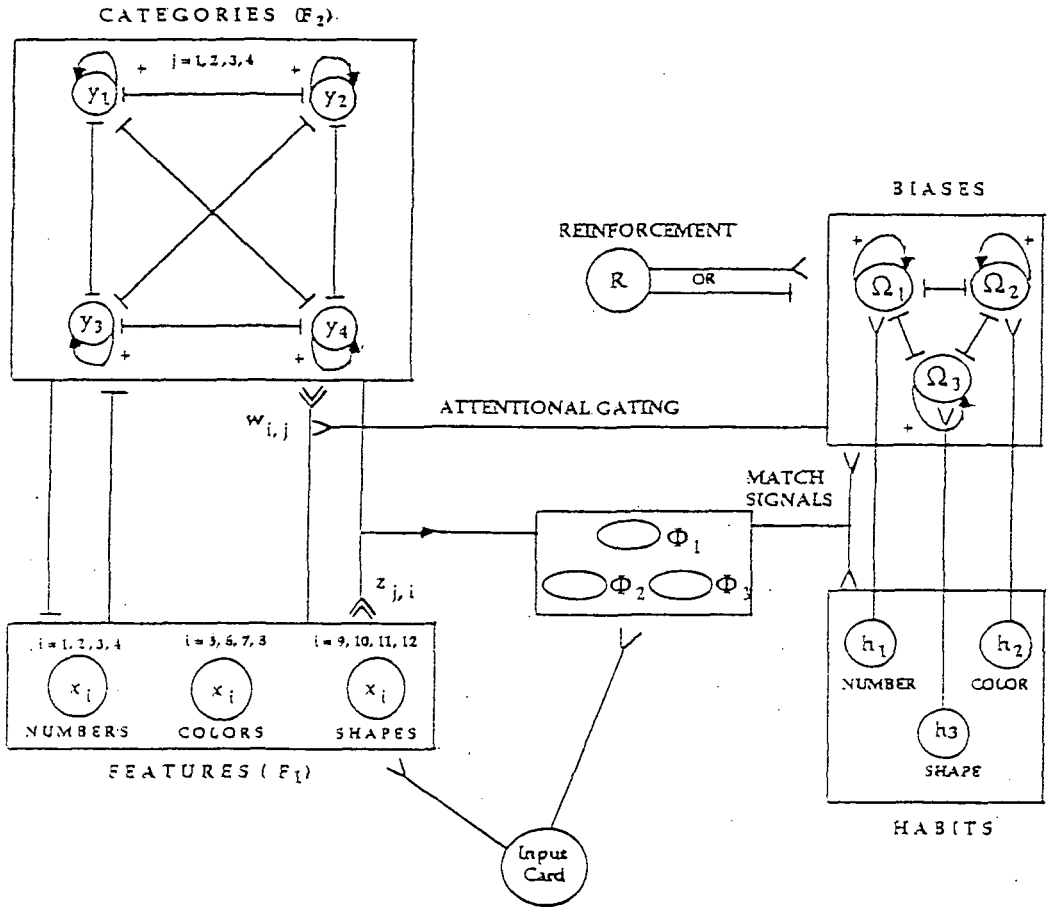


Figure 3: The network developed by Leven and Levine, based on ART, to simulate Milner's card sorting data. The reinforcement signal R is applied to the bias nodes Ω. The bias nodes in turn gate the category neurons. The bias neurons also get input from the habit neuron, which encodes past categorizations. The match signals (which state the particular criterion of the input that was responsible for categorization) are encoded by Φ nodes. The feature neurons x_i encode the features of the input card and categorization is achieved depending on previous reinforcements

$$= \quad -Dy_j(\sum_{r \neq j} f(y_r) + \mathcal{I}), \qquad j = 1,2,3,4, \tag{3}$$

One of the excitatory term in the above equation, $g(\Omega_{[\frac{i+3}{4}]}x_i)w_{i,j}$, consists of feature node activities gated by bias nodes and bottom up weights. The other excitatory term is the positive self feedback of the node itself. The inhibitory term consists of mutual lateral inhibition given by "$\sum_{r \neq j} f(y_r)$" and the reset signal "$\mathcal{I}$". The function "g" in the above differential equation is as follows.

$$\begin{aligned} g(x) \quad = \quad & 0, & x < 0.5 \\ & x - 0.5, & 0.5 \leq x \leq 3 \\ & 2.5, & x > 3. \end{aligned} \tag{4}$$

When a "response card" is input to the network, the feature neurons that encode the characteristics of that input become active. These active feature neurons excite the respective category neurons. Consider the example of two red crosses input "response card" again. This card could excite "category 1" (because of color), "category 2" (because of number) or "category 3" (because of shape). Now, suppose that categorization with respect to color has been rewarding in the previous attempts. Then, the habit and bias nodes encoding color feature would be more active than the other criterions. This would lead to a greater activity of the second category neuron. The faster than linear function for the self-feedback and lateral inhibition terms causes the second category neuron to become more active and also, at the same time, suppress the activity of the other category neurons[2]. Thus, the input "response card" would be matched to the first "stimulus card" category since the previously rewarding criterion had been color.

The habit and bias network are comprised of three neurons, one for each criterion: number, color and shape. The habit nodes detect how often a dimension (number, color or shape) is used to determine categorization of the "response card" regardless of whether that categorization is rewarded or punished. The activity of the habit nodes is given as follows

$$\frac{dh_k}{dt} \quad = \quad Hh_k\{(J - h_k)[\Phi_k - \theta_2]^+ - [\Phi_k - \theta_2]^-\} \qquad k = 1,2,3 \tag{5}$$

where θ_2, H, and J are positive constants. The functions $[x]^+$ and $[x]^-$ used in the above differential equation imply the following

$$\begin{aligned} [x]^+ \quad = \quad & x & & if \quad x > 0 \\ = \quad & 0 & & else \end{aligned} \tag{6}$$

[2]For a detailed study of different type of functions i.e slower than linear, linear and greater than linear and the roles they play in lateral inhibition and self-excitation the reader is referred to (Grossberg, 1973).

$$
\begin{aligned}
[x]^- &= -x & \quad if \quad x < 0 \\
&= 0 & \quad else
\end{aligned} \tag{7}
$$

Φ_k is called the "match signal", and is given by

$$
\Phi_k = \sum_{i=4k-3}^{4k} z_{j,i} I_i, \qquad k = 1,2,3 \tag{8}
$$

where j is the index of the chosen category, $z_{j,i}$ is the top-down weights and I_i is the input vector of the "response card". For example, if the input card is two red crosses (i.e., 0100 1000 0010) and sorting based on color had been rewarding, then the input card would be sorted by color. For the habit node representing color, k = 2. Thus, the range of the summation for calculating the match signal goes from I_5 to I_8. Now I_i is positive for i=5 (the color red) and $I_i = 0$ for i= 6,7,8. Hence, the top down weights of only the "stimulus card" category chosen (i.e j=1, which represents one red triangle) would be gated with the input signals I_5 to I_8 to form the match signal for the number match node. This ensures that only the criterion that was used to sort the input card is credited with the successful match. Thus only the "match signal" containing the information regarding color becomes much larger than the other match nodes. Hence the activity of the habit node for which the match signal is large would increase and the activity of the rest of the habit nodes would decrease. The activity of a habit node for a particular dimension corresponds to the past frequency with which that dimension was used to categorize the input "response card".

The Bias nodes, on the other hand, encode the recent past success of using the corresponding dimension (number, color, shape) to categorize the input "response card". As such, the Bias nodes receive direct input from the reinforcement signal. A positive reinforcement signal (R^+) implies that the criterion used to sort the input "response card" is correct and a negative reinforcement signal (R^-) implies the opposite. The activities of the bias nodes are given below.

$$
\begin{aligned}
\frac{d\Omega_k}{dt} = & -E\Omega_k + \{(F - \Omega_k)([h_k - \theta_1]^+ + \alpha R^+ + g(\Omega_k)) \\
& -\Omega_k(\alpha R^- + G\sum_{r \neq k} g(\Omega_r))\} f(\Phi_k) \qquad k = 1,2,3
\end{aligned} \tag{9}
$$

In the above equation both the excitatory and the inhibitory terms are gated by the match signals so that the reinforcement signals (either positive or negative) are conveyed to the appropriate bias neuron. Like the habit nodes there are three bias neurons. The excitatory term in the above equation, that is $(F - \Omega_k)([h_k - \theta_1]^+ + \alpha R^+ + g(\Omega_k))$, comprises of an external reward signal R^+ multiplied by a gain factor α, the habit node activity and a positive self feedback term $g(\Omega_k)$. The inhibitory term consists of $\Omega_k(\alpha R^- + G\sum_{r \neq k} g(\Omega_r))$ where R^- is the external punishment signal multiplied by the gain α and

the activities of the rest of the bias nodes. Coming back to our example (i.e the two red crosses) suppose that the sorting of the input "response card" according to color was correct and the reward signal (R^+) was briefly turned high. [3] This would increase the activity activity of the color bias node which would encourage future sortings of input "response card" by color. The increase in the activity of "color" bias node would cause the decrease in the activity of the other two bias nodes due to inhibition (thus indirectly leading to an increase in its own activity). On the other hand, supposing the sorting criterion used to categorize the input "response card" was inappropriate, then, a punishment signal would be given which would cause a decrease in the activity of the respective (in this case the "color") bias node. This would lead to an increase in the activities of the other two bias nodes. Eventually categorization of future input "response cards" would be by a criterion other than "color".

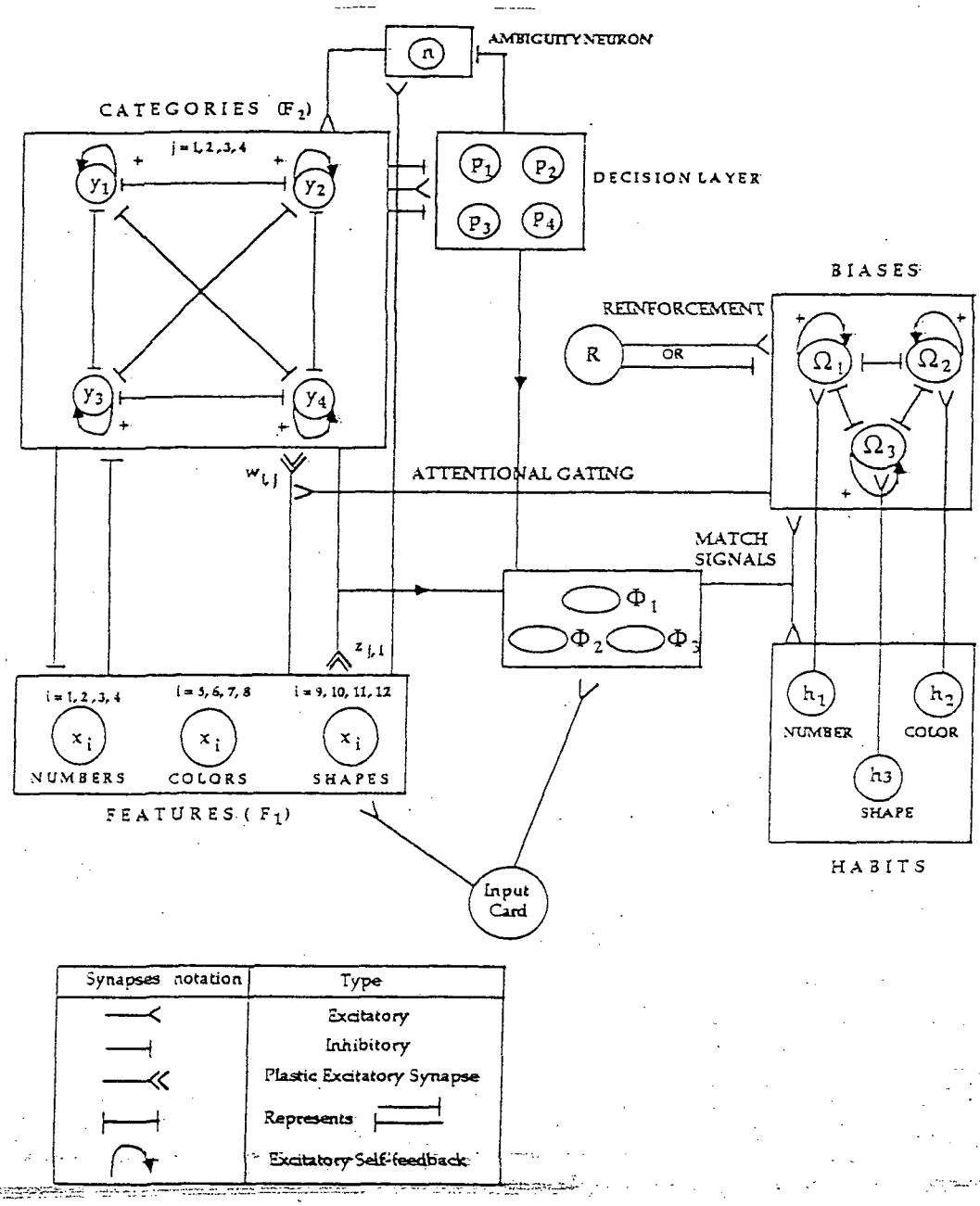
The gain term α that multiples the reward and punishment signals represents the influence external signals have on the categorization of input "response cards". Normal subjects capable of changing their criterion of sorting according to external reinforcement signals have a high value for α. Those who are incapable of this task due to lesions in their frontal lobes have a low value for α. Thus the above network illustrates how the criterion for categorization is dependent on external reinforcement signal.

3.4 Modifications

The model proposed by Leven and Levine(Leven and Levine, 1987) as discussed above comprises of two parts: the Categorization network and the Habit and Bias network. The simulations that they conducted using the above differential equations also were conducted in a two stage manner. In the first stage the input vector was presented to the Categorization network which sorted the input "response card" to a specific category depending on which of the criterions had been rewarding. Once the input had been sorted to a particular category the match signal was calculated algorithmically for the given input. Then, in the second stage, the Habit and Bias node activities were simulated along with the reinforcement signal.

In order to obtain a continuous-time non-algorithmic model, we replaced the algebraic match signal by dynamic match signal nodes. Moreover, in order to make the system self-aware as to when a categorization choice was made, a decision layer was incorporated. A cognitive node was also designed to solve the ambiguity in case a decision regarding the category was not made.

[3] The reward and punishment signals R^+ and R^- are large positive and negative pulses which cause a fast "rise" of the bias node response. The decay time constant of the bias and habit nodes are much smaller than the time constants of the feature and category neurons. Hence the activity of the bias and habit nodes tend to change much slower than the activities of the feature and category nodes during categorization of input "response cards".



Synapses notation	Type
	Excitatory
	Inhibitory
	Plastic Excitatory Synapse
	Represents
	Excitatory Self-feedback

Figure 4: The modified version of the Leven and Levine model consists of an additional decision layer and an ambiguity neuron. The decision layer ensures that only one of the category is selected. The ambiguity neuron resolves any indecisions that occur in the network as to how an "input card" should be sorted. This indecision is resolved randomly. The match signals are generated when a particular card has been categorized according to a particular criterion. The bias nodes modulate the categorization according to reinforcement signals and habits.

The modified neural network is shown in Figure 4. The dynamic signals formerly computed algorithmically in the original Leven and Levine (Leven and Levine, 1987) are computed in continuous-time in the modified model. The match signal node activity is given below.

$$\begin{aligned} \frac{d\Phi_k}{dt} &= -A\Phi_k + (B - C\Phi_k)\{\sum_{j=4k-3}^{4k}\sum_{i=1}^{4} I_j g_1(p_i - \theta_1) z_{i,j}\} \\ &\quad -D\Phi_k\mathcal{I} \qquad k = 1,2,3, \end{aligned} \tag{10}$$

A, B, C, θ_1 and D are positive constants. The match signal comprises of an excitation term $(B - C\Phi_k)\{\sum_{j=4k-3}^{4k} I_j g_1(p_i - \theta_1) z_{i,j}\}$ where I_j is the input vector, p_i is the activity of the neurons in the decision layer, θ_1 is a constant and $z_{i,j}$ is the top-down weights from the category nodes to the feature nodes. The inhibition term given by $-D\Phi_k\mathcal{I}$ consists of a reset signal $\mathcal{I}$ (c.f eqn 3) and a summation of the other match signals as input to a "linear above threshold function" g_1. The function f is the same "faster than linear function" used in original model.

Consider the example discussed previously (in light of the new modifications) in which the input "response card" was two red crosses. That input is now given to the new network. Assume that sorting by the "color" criterion had been rewarding. Then the input card would be categorized as being in the first category (which corresponds to one red triangle). However, during sorting, various category nodes might have varying activities. To generate the correct match signal the categorization must be completed. The decision layer which has been included into the modified network ensures this.

The decision layer, comprised of four nodes (one for each "stimulus card" category), has feed-forward connections from the category nodes. The following is the differential equation for one such decision node.

$$\frac{dp_i}{dt} = -A_1 p_i + (B_1 - p_i) W y_i - p_i W \sum_{j \neq i} y_j \qquad i = 1,2,3,4, \tag{11}$$

where W represent a constant gain by which the inputs to the decision layer are multiplied. This equation represents a shunting on-center off-surround type feed-forward network. p_i encodes the ratio of the y_i node activity to the total activity of the categorization layer. By setting a threshold level for the activity of the decision layer neurons it can be determined if categorization is achieved. When categorization has been achieved, only one of the decision making neurons is active and that neuron represents the category to which the input "response card" has been sorted.

Consider the case when no previous categorization has been done. When an input card (say two red crosses) is given to the categorization network, no biases for sorting the input card according to any specific criterion (i.e number, color or shape) exist since neither of the criterions is more preferred than

the other. However, in case of humans, since a decision has to be made, one arbitrarily selects one of the three category nodes for the card (forced choice paradigm). This higher cognitive decision making process is modeled by an ambiguity neuron "n" whose differential equation is given by

$$\frac{dn}{dt} = \Gamma\{-An + (B-n)\{\sum_{i=1}^{12} I_i\} - n\{\Upsilon \sum_{j=1}^{4} g_1(p_j - \theta_1) + \mathcal{I}\}\} \tag{12}$$

The excitation $(B-n)\{\sum_i I_i\}$ term consists of summation of all input features I_i so as to make sure that this neuron is active only when an input card is presented to the categorization network. The inhibition term consists of a high gain parameter Υ which multiplies the summation of the activities of the decision layer nodes. This is to ensure that if categorization of the input card has been achieved then the ambiguity neuron activity should be inhibited. The inhibition term also consists of a reset signal $\mathcal{I}$ which resets the ambiguity neuron before a new card is input to the categorization network. Γ When the categorization network is incapable of sorting the input card to a particular category, the ambiguity neuron introduces a random bias to the inputs of the category neurons which leads to the categorization of the input card to one of the category nodes. To incorporate this random bias, the activity of the category neurons is modified to the following differential equation.

$$\begin{aligned} \frac{dy_j}{dt} &= -Ay_j + (B - Cy_j)(f(y_j)(1 + [n-\theta_2]^+ v_j{}^n) + \sum_{i=1}^{12} g(\Omega_{[\frac{i+3}{4}]} x_i) z_{i,j}) \\ &\quad -Dy_j(\sum_{r \neq j} f(y_r) + \mathcal{I}), \qquad j = 1,2,3,4 \end{aligned} \tag{13}$$

The above differential equation includes only an additional excitatory term $(1 + [n-\theta_2]^+ v_j{}^n)$ in comparison to the differential equation for the category neuron in the original model discussed in the previous section. In the above equation $v_j{}^n$ is the random connection weight between the ambiguity neurons and the respective category nodes and θ_2 is a threshold constant.

The differential equations for feature neurons, for the habit neurons and the bias neurons were similar to those used in the original model described in the previous section. This concludes the modifications done to the Leven and Levine's model.

3.5 Simulation Results

3.5.1 ART network

An ART-1 network designed by Carpenter and Grossberg was simulated [4] to categorize alphabets(Carpenter and Grossberg, 1987). The ART-1 network is shown in Figure 5 and it comprises of two layers, the F_1 and F_2 which encode patterns of activation in short-term memory (STM). The bottom-up and top-down pathways between F_1 and F_2 contain adaptive long-term memory (LTM) traces. Various handwritten alphabets were captured using a vision system and converted to binary images. These alphabets were then used as inputs to the ART-1 network and the various categorizations were observed for different vigilance parameters. Figure 6 shows such categorization of alphabets for a vigilance coefficient of 0.95 and 0.93. The left hand side shows the set of images used to train the network. The figure on the right shows the top down weights of the category into which the input was categorized for vigilance parameter 0.95 and 0.93.

3.5.2 Reinforcement-habit network

Milner(Milner, 1963, 1964) using a modified Wisconsin card sorting test, showed that people with frontal damage were incapable of shifting their decision making criterion dynamically in presence of reward and punishment. Leven and Levine(Leven and Levine, 1987) modeled this scenario using a neural network which was capable of replicating the behavior of both frontally damaged and normal human depending on the gain constant α. Figure 4 shows the modified version of the Leven-Levine model. As discussed previously the modified model comprises of a decision layer which ascertains whether the category neurons have come up with a decision regarding the category to which the input belongs. Figure 7 demonstrates the case when a category choice is made for a given input. In this simulation the input consisted of two red crosses (i.e 0100 1000 0010). The categorization in the previous cases done according

[4] All the simulations were initially conducted on an Ultrix Workstation. The X-window environment was used for the interactive graphics in the simulations. All the code was written in C and the operating system used was Unix. The simulations were conducted in a continuous-mode so as to model real life situations. The differential equations were solved using Runge-Kutta-Falsberg algorithm which was obtained from the Oak-Ridge lab. These simulations were then ported to the Sparc Station at NASA-JSC on which they are running. Parallel implementation of these simulation was also conducted on the Amdahl at NASA-JSC. Each of the simulations consist of a pair of window called the input and the output windows. They represent respectively the input to and the response of the modeled neural network. Other windows represent the activity of the various neurons and the strengths of the signals. The simulation once evoked begins running continuously and can be interrupted any time by typing Ctrl-C. On doing this a menu pops up listing the various options that can be chosen. The various options range from input of a new object to the network, to quitting of the entire simulations.

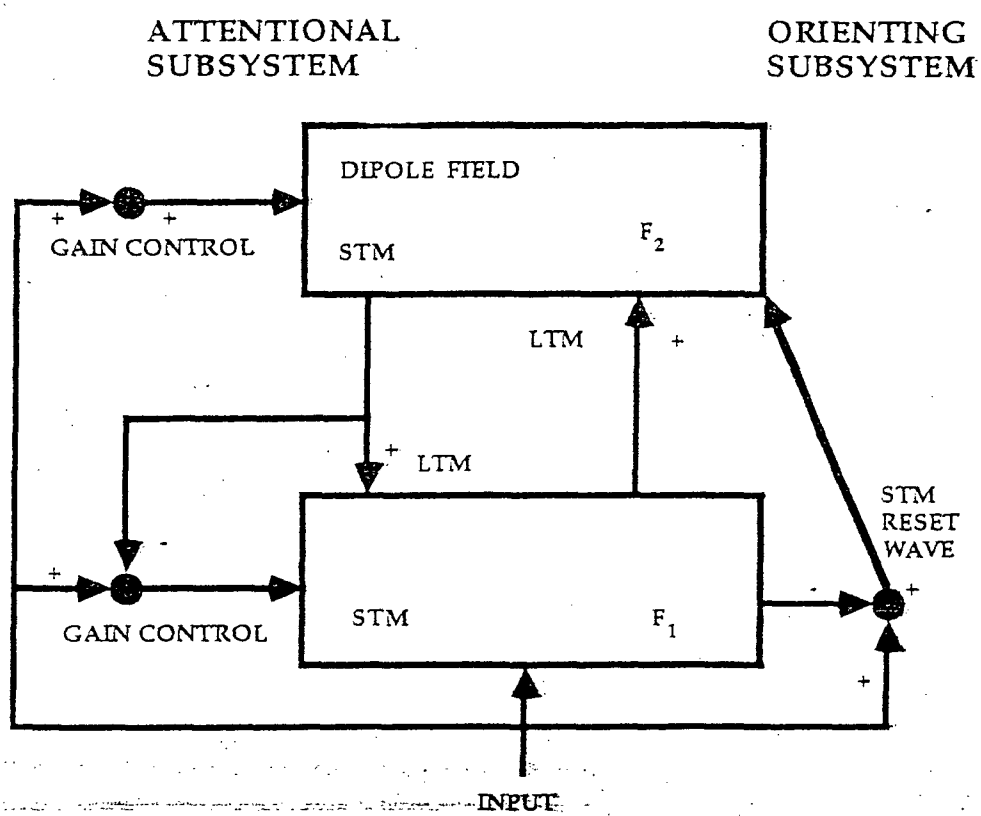
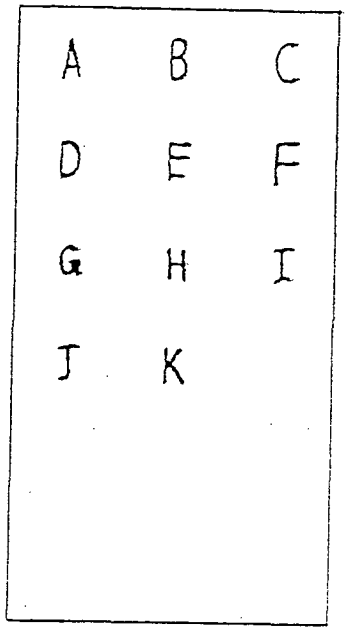
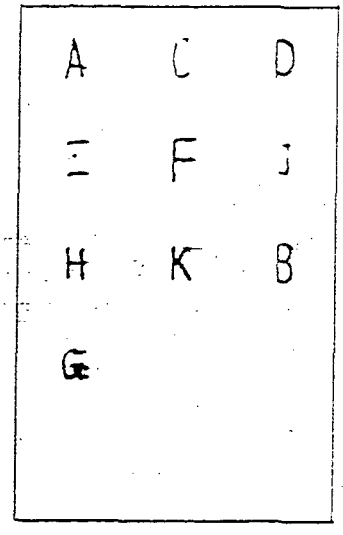


Figure 5: The Adaptive Resonance Theory model comprises of two layers, the F_1 layer is the input layer and F_2 is the categorization layer.

Inputs



Top-down weights (vigilance 0.95)
1. A 2. C 3. D 4. E, I 5. F 6. J 7. H
8. K 9. B 10. G



Top-down weights (vigilance 0.93)
1. C 2. B, E 3. J, I 4. A 5. G, H 6. F
7. D 8. K

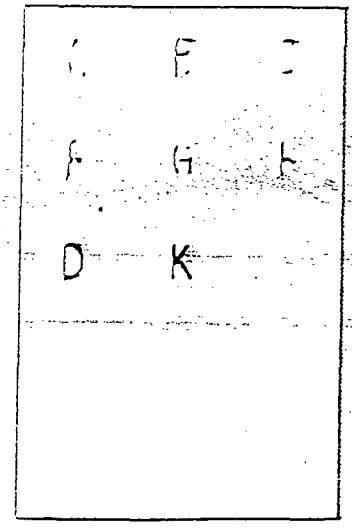


Figure 6: The input comprises of 8 alphabets from "a" to "k" and the output gives the top down map of the classes into which the various inputs were categorized for vigilance parameter 0.95 and 0.93.

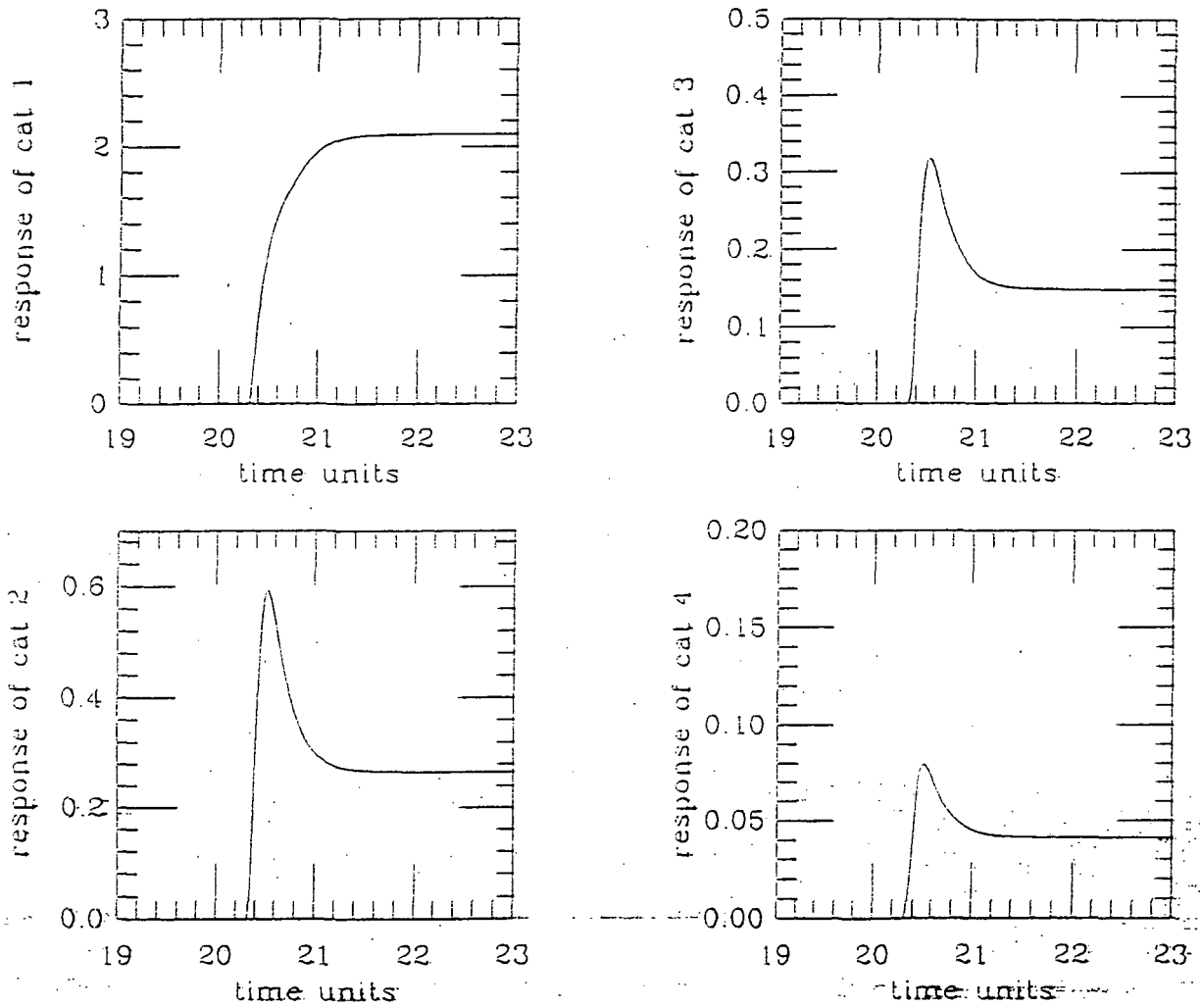


Figure 7: The four graphs show the response of the four respective category nodes. The results shown are for a simulation in which the input "response card" to the network was 2 red crosses. Since the previous categorizations had been by color, the category neuron "1" representing one red triangle has the maximum activity. Thus the input response card has been categorized "correctly" by color.

to color had been rewarding and in this case the network should categorize according to color. The graph shown in figure 7 consists of the activity of the four category neurons. As can be seen from the graphs, the category neuron "1" has the highest activity (which represents the one red triangle) implying that the input card was matched according to color.

The second simulation (figure 8) shows the situation in which categorization was made possible by the ambiguity neuron. The response card two red crosses is the input to the model. Initially categorization of inputs was not possible because there was no information about any criterion being more rewarding than the other. However, the initial indecision is resolved randomly by an ambiguity neuron. As can be seen from the graph, the second decision neuron activity rises above the threshold implying that the model matched the input to the second category.

4 Novelty and Reinforcement

4.1 Neurophysiological basis

Novelty is yet another facet in our decision making process. In general new objects attract human attention. However, such a new (novel) stimulus looses its attraction once it has been around long enough. A baby who is shown a new toy is quickly attracted to it and many times leaves what she is holding or doing to acquire the new toy. Thus, the new toy (which is novel) is more attractive to her than the other things already existing within her reach (environment). Now, if one were to show her a new toy after a sufficient time, she would leave the former toy for the later one, as the former toy would have lost its novelty[5]. Novelty plays a important role in the exploratory behavior of humans and is vital for learning new things.

Experiments conducted by Pribram on normal and frontally lesioned rhesus monkeys suggested that the frontal lobes play a major role in novelty based behavior(Pribram, 1961). One such experiment consisted of a scene comprising of a board having twelve holes. Two junk objects were placed on two of these holes and under one was placed a peanut. The task of the monkey was to open the object under which the peanut was placed (a rewarding event). Once any one of the objects had been lifted by the monkey then the wooden board was hid from the monkey's view and another peanut was replaced under the object if the monkey had picked the object containing the peanut. After a certain number of successful trials in which the monkey picked the object under which the peanut was kept, the reward (i.e the peanut) was moved to the second object. This process was continued until the monkey had completed a fixed number of successful trials. After this a *new* i.e a third junk object, was placed in

[5]Piaget's treatise on cognitive behavior in children (Elkind and Flavell, 1969)

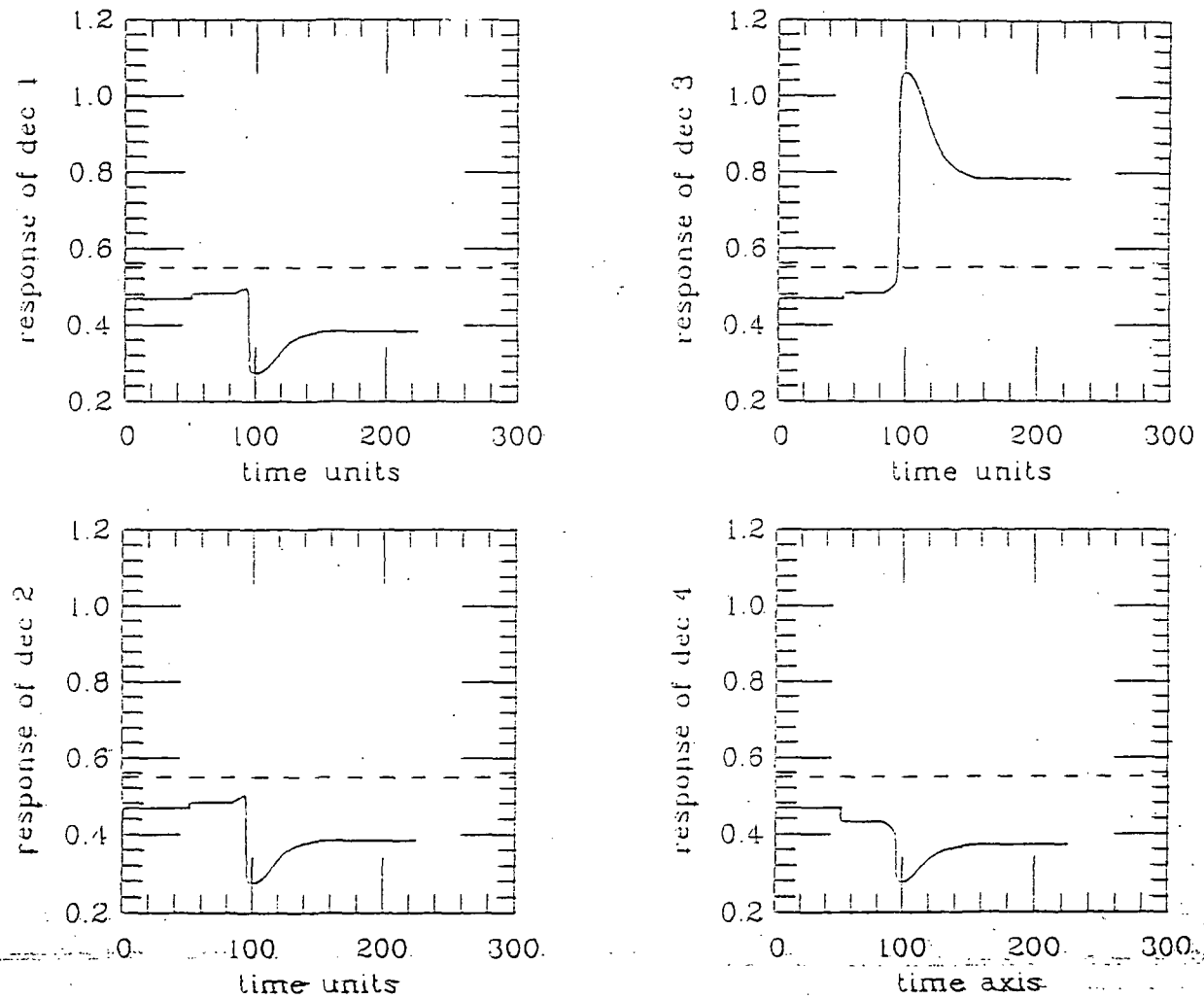


Figure 8: The four graphs shown in this figure represent the activities of the four decision layer neurons. When an response card was initially input to the ART network it was not able to categorize it. Eventually categorization was randomly achieved as the input card was categorized to category "3".

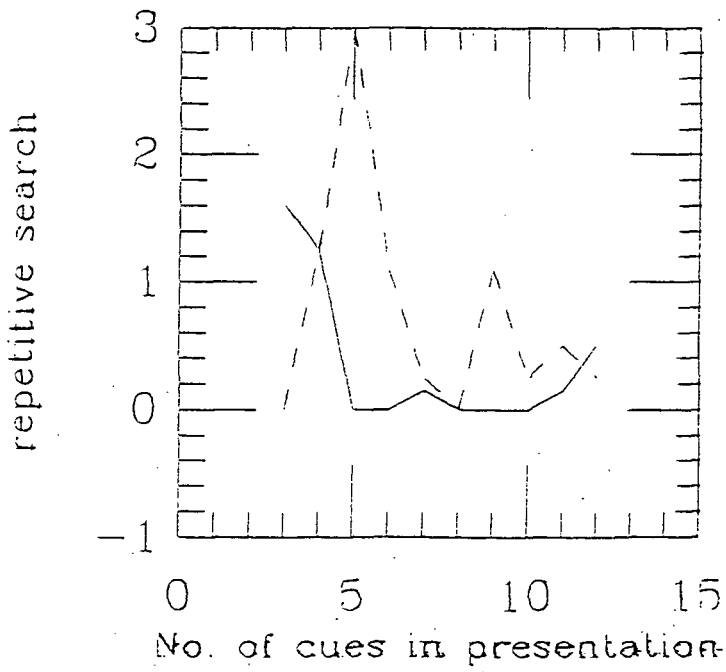


Figure 9: The experimental data shown above represents the number of repetitive trials required by frontal monkeys (straight lines) and normal monkeys (dashed lines) to pick up a novel object in Pribram's experiment.

the scene. Under this new object was kept the peanut. Thereafter the number of junk objects were increased after a fixed number of successful picks by the monkey, until all the 12 wooden holes were covered.

Figure 9 shows the results of Pribram's experiment. The number of repetitive errors (lifting of previously rewarding objects) until it picked the new object containing the reward is plotted against the number of junk objects. In general it can be seen that frontally damaged monkeys make fewer errors than normals, being more attracted to novel objects[6]. Thus, Pribram concluded that in normal monkeys the frontal cortex was responsible for suppression of novelty driven behavior. As in the rhesus monkey, the human the frontal lobes also seem to be responsible for gating novelty based cognitive decision making.

4.2 Significance

Attraction to novelty is essential for exploration and learning new things. However, extreme attraction to novelty in a changing environment can be disastrous and should be weighted by reinforcement signals. Pribram's experiment show that if this balance is disturbed (by a damage to frontal lobes that would weaken the effect of reinforcement signals), pathological behavior results. Again, comparison of pathological and normal subjects reveal "transparent" aspect in decision making.

4.3 Neural models

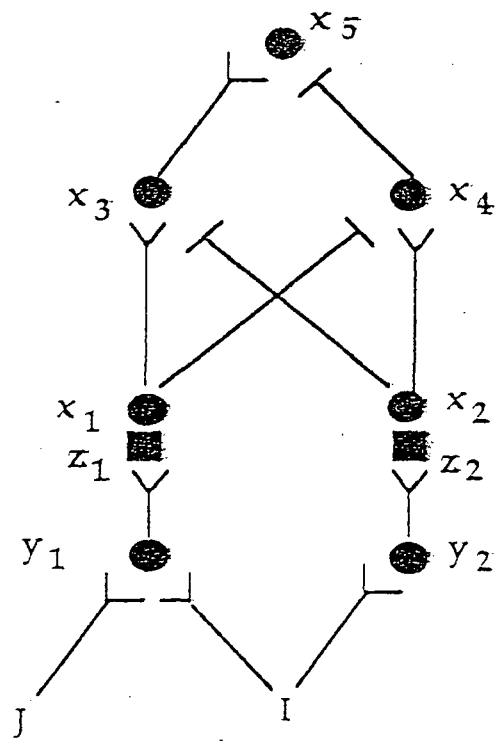
Levine and Prueitt(Levine and Prueitt, 1989) modeled Pribram's psychological findings with a neural network comprising of gated dipoles. Gated dipoles are devices using transmitter mechanisms for comparing present and past signal values. An example of a gated dipole is shown in figure 10. The square synapses in this figure signify such transmitter packages and their activities are given by the following equations.

$$\frac{dz_1}{dt} = a_1(\beta - z_1) - a_2 y_1 z_1 \tag{14}$$

$$\frac{dz_2}{dt} = a_1(\beta - z_2) - a_2 y_2 z_2 \tag{15}$$

where z_1 and z_2 are the amounts of available transmitter at the ON and OFF channels respectively, a_1 is the rate of accumulation of the transmitter to the maximum value of β (the total amount of

[6]Details regarding improved performance of normal monkeys in case of larger number of objects and worse performance of frontally damaged monkeys in case of few objects have been discussed by Levine and Prueitt & Pribram(Levine and Prueitt 1989; Pribram, 1961)



Synapses notation	Type
	Excitatory
	Inhibitory
	Transmitter Gated Synapse

Figure 10: The schematic diagram of a gated dipole

transmitter) and a_2 represent the depletion rate of the transmitter. The activities of the input node y_1 and y_2 are given by

$$\frac{dy_1}{dt} = -gy_1 + I + J \tag{16}$$

$$\frac{dy_2}{dt} = -gy_2 + I \tag{17}$$

where g is the decay rate, I is the background arousal signal and J is the input. The activities of the remaining nodes x_1, x_2, x_3 and x_4 are given by

$$\frac{dx_1}{dt} = -gx_1 + by_1z_1 \tag{18}$$

$$\frac{dx_2}{dt} = -gx_2 + by_2z_2 \tag{19}$$

$$\frac{dx_3}{dt} = -gx_3 + b[x_1 - x_2]^+ \tag{20}$$

$$\frac{dx_4}{dt} = -gx_4 + b[x_2 - x_1]^+ \tag{21}$$

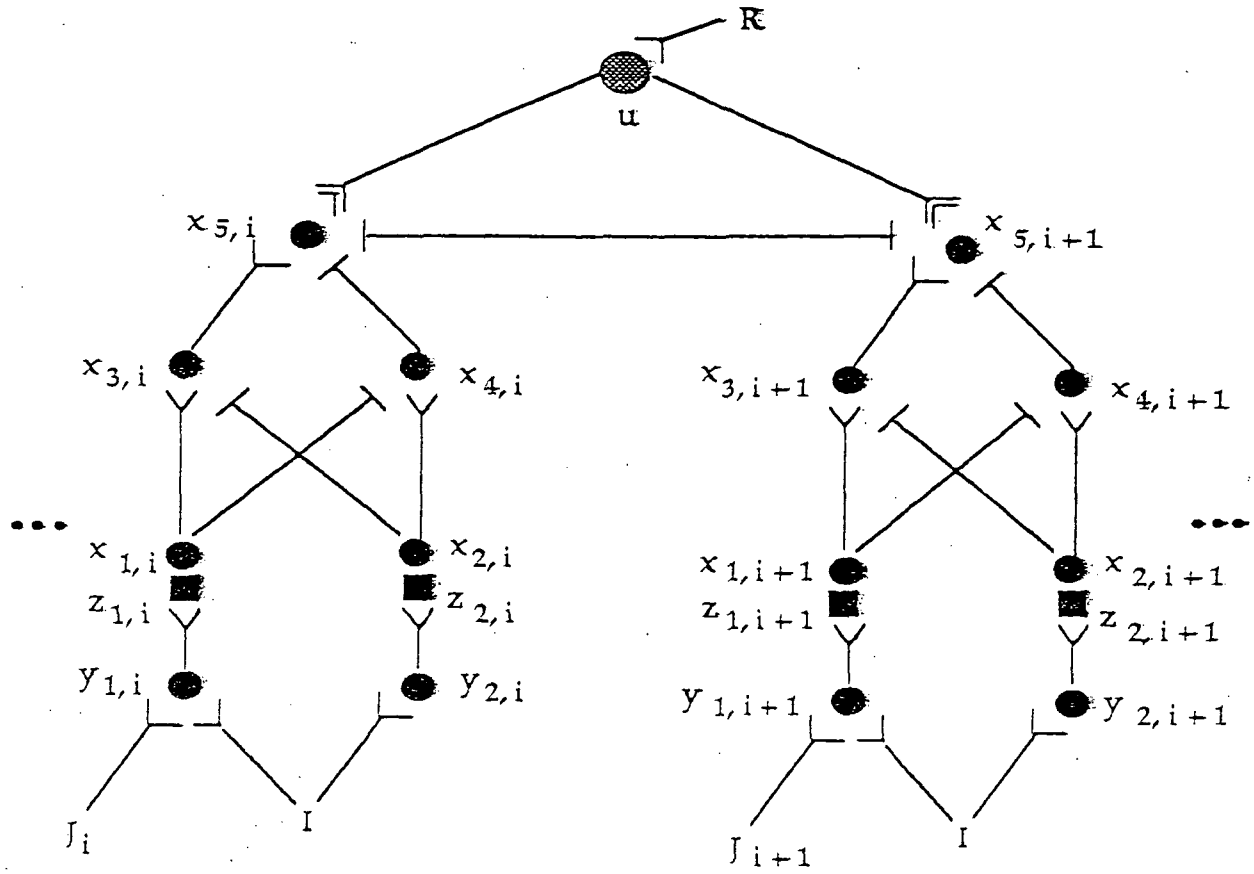
$$\frac{dx_5}{dt} = -gx_5 + (1 - x_5)x_3 - x_5x_4 \tag{22}$$

where b is a constant.

Now consider when, no input is applied. The amounts of transmitter z_1 and z_2 are identical, hence x_5 neuron is not active. When the input J is turned "on", the depletion of transmitter in the left channel is more than that in the right channel. Hence z_1 is less than z_2. However the greater input (I + J, in the left channel) overcomes the more depleted transmitter to make the left channel more active. When the input J is turned off the right channel becomes transiently active and inhibits the activity of the x_5 neuron until the transmitter of the left hand synapse is completely replenished. This causes the rebound of the x_5 activity. [7].

Figure 11 shows the neural network proposed by Levine and Prueitt to explain Pribram's psychological results. This network comprises of a number of gated dipoles coupled to each other (only two gated dipoles are shown for simplicity and the subscript i implies a specific gated dipole). Each of the gated dipole corresponds to an object in the environment. The presence of an object in the environment is indicated by the input J ("on" implies presence of an object and "off" implies its absence). The activity of $x_{5,i}$ neurons indicate the possibility that the model is chosing that particular object. Thus the $x_{5,i}$ node with the highest activity is the object that the model has chosen. A nonspecific arousal input I is

[7] For a detailed discussion of the gated-dipole dynamics refer to Grossberg(Grossberg, 1972)



Synapses notation	Type
	Excitatory
	Inhibitory
	Transmitter Gated Synapse
	Plastic Excitatory Synapse
	Plastic Inhibitory Synapse
	Represents

Figure 11: The original model proposed by Levine and Prueitt.

also given to all the gated dipoles in the network. The activity of the network is modulated by reward signal via the reward node whose differential equation is

$$\frac{du}{dt} = -gu + r \tag{23}$$

where u is the reward node activity and r is the reward signal. This reward node modulates the activity of the gated dipole node $x_{5,i}$ as follows

$$\frac{dx_{5,i}}{dt} = -gx_{5,i} + (1 - x_{5,i})(euw_{i,5} + x_{3,i}) - cx_{5,i}(x_{4,i} + \sum_{j \neq i} x_{5,j}) \tag{24}$$

where one of the excitatory terms consists of weighted activity of the reward node multipled by a coupling parameter e (between the reward node and the gated dipole node $x_{5,i}$) and the other term is the activity of its excitatory neuron $x_{3,i}$. The inhibitory term consists of two terms: $\sum_{j \neq i} x_{5,j}$ which constitutes the lateral inhibition of the other gated dipoles and the inhibitory $x_{4,i}$ node of the gated dipole. The behavior of normal monkeys is modeled by a large value of the coupling factor "e" and a small value of the coupling factor models the behavior of frontally lesioned monkeys. In case of normal monkeys the reward signals contribute significantly to the activity of the respective $x_{5,i}$ neuron. The effects of the reward signals are stored in long term memory via modification of the weights connecting the reward node and the respective $x_{5,i}$ neuron. The coding of reward signals into weights is given by

$$\frac{dw_{5,i}}{dt} = -f_1 w_{5,i} + f_2 u x_{5,i} \tag{25}$$

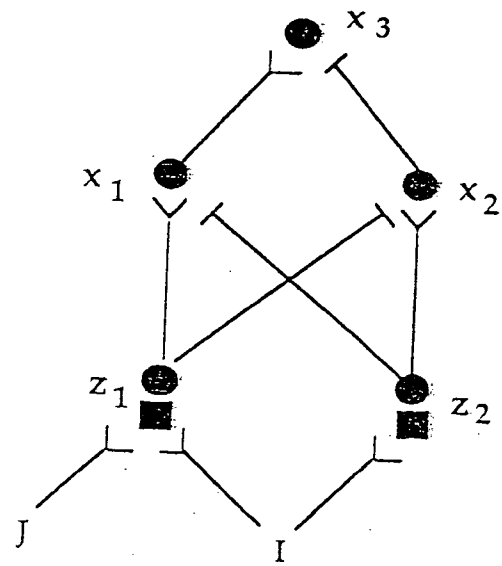
where, $w_{5,i}$ is the synaptic strength between the reward node u and $x_{5,i}$, f_1 and f_2 are positive constants.

Thus for low values of the coupling parameter e novelty of new objects plays the major role; however for large values of e previously rewarding objects can override the attraction of novel objects.

4.4 Modifications

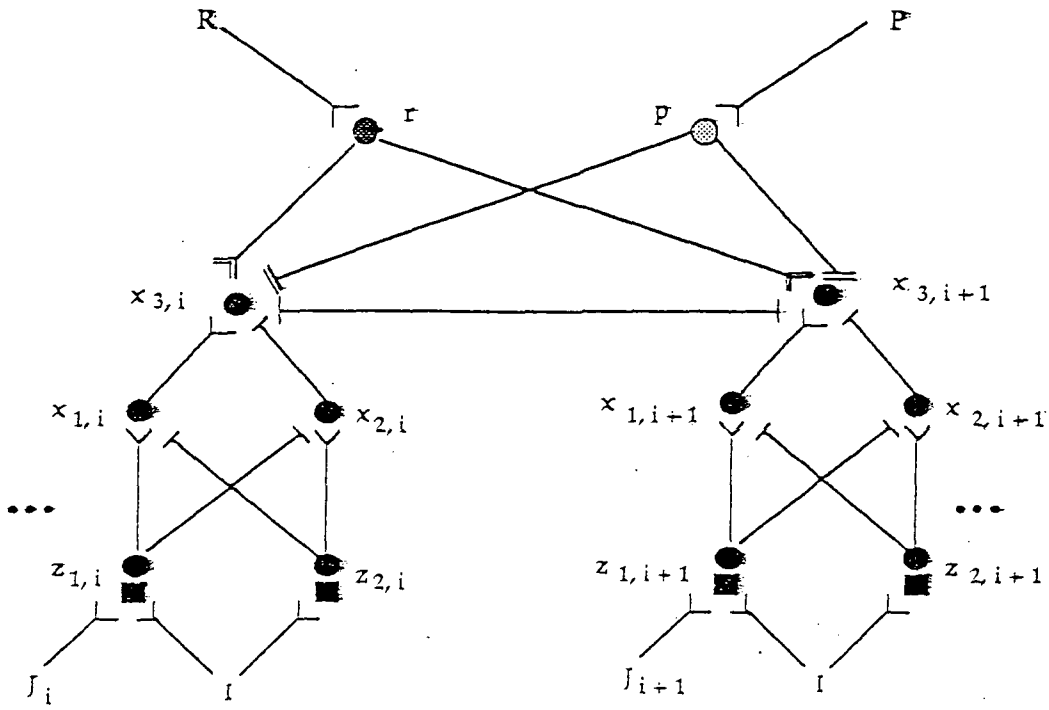
The above model was modified and used in the cognitive unit of the self-organizing robot. The original gated-dipole network was replaced by a simpler version shown in Figure 12. The "ON" channel of this gated-dipole comprises of the transmitter node z_1 and the node x_1. The "OFF" channel comprises of the transmitter node z_2 and the node x_2. The gated-dipole node x_3 (similar to the x_5 node in the original gated-dipole model) combines the inputs of both the "ON" and "OFF" channels. Using these modified (simple) gated-dipole network a modified version of the Levine-Prueitt model was developed. This modified model is shown in Figure 13.

In the Levine-Prueitt model although the effect of a reward could decay through the extinction mechanism in learning, there were no mechanisms to accommodate punishment effects. Though a negative



Synapses notation	Type
—≺	Excitatory
—⊣	Inhibitory
■	Transmitter Gated Synapse

Figure 12: Schematic of the modified gated dipole network.



Synapses notation	Type
	Excitatory
	Inhibitory
	Transmitter Gated Synapse
	Plastic Excitatory Synapse
	Plastic Inhibitory Synapse
	Represents

Figure 13: The schematic representation of the modified Levine-Prueitt model.

reward input may seem to enable the model to incorporate punishment by using the same connection, a closer examination reveals that such a strategy would create an ambiguity due to the asymmetry of reward and punishment signals. Assume that a synaptic weight has been increased due to a repetitive application of reward signals for a particular object. Such a reinforcement would indirectly increase the effect of punishment input as well. The single LTM scheme cannot independently store in LTM reward and punishment signals and thus creates an ambiguity. In order to achieve an independent LTM storage, we introduce separate LTM traces using separate reward and punishment synapses. However, this solution implies a subtle asymmetry property for encoding reward and punishment by Hebbian synapses. A reward signal causes an increase in the post-synaptic neuron activity and thus the LTM trace correlates increases in pre and post synaptic activities and reinforces this increase. An asymmetry arises in punishment: A punishment signal depresses the post-synaptic activity. The LTM trace of punishment correlates increasing pre-synaptic activity with decreasing post-synaptic activity. In a Hebbian synapse this asymmetry causes a major problem since the inhibiting synapse cannot effectively code the STM traces. Moreover such a coding would also be relatively insensitive to punishment with respect to reward (which is not desirable). To solve this problem, we introduced a new type of learning rule. Both the reward and punishment nodes have the same dynamics:

$$\frac{dr}{dt} = -A_1 r + (B_1 - r)R + C_1 \tag{26}$$

$$\frac{dp}{dt} = -A_1 p + (B_1 - p)P + C_1 \tag{27}$$

where A_1, B_1 and C_1 are constants and R and P are external reward and punishment signals respectively. The reward and punishment node dynamics have been modified to shunting type equations. Shunting equations were used so as to be able to dissociate learning rates from forgetting rates(Grossberg, 1988). While the learning equations for the reward weights remain the same the learning equations for punishment weights has been modified to include an auxiliary variable y_i. The activity of this variable is as follows.

$$\frac{dy_i}{dt} = -A_3 y_i + (B_3 - y_i)\gamma_1 g(x_{3,i} - \theta) \tag{28}$$

The synaptic modification equation is given by:

$$\frac{dw_{p,i}}{dt} = -A_2(w_{p,i} - 1) + (M_p - w_{p,i})B_2 g(p - \theta_1)g(y_i - \theta_2) \tag{29}$$

where γ_1, A_3, A_2, B_3, B_2 θ, θ_1, θ_2 and M_r are constant and g is a linear above threshold function. y_i reflects past values of postsynaptic neuron to enable the association of the current punishment signal with past activities of decisions.

Another problem of the original model was that an external monitor (experimenter) was needed to ascertain whether the network chose a particular object from the environment. Such an implementation in a continuous environment is not possible and hence an additional network capable of handling this situation was implemented using a feed-forward shunting network as shown in figure 14. This network receives inputs from the gated dipole circuit and when only one of the nodes is active, sends a signal to the robotic arm (This network is similar to the decision layer discussed in the modified Leven-Levine model).

The differential equation of the gated dipole i.e for x_3 (the node x_5 in case of original model) was also modified to incorporate the various changes. The modified equation is as follows.

$$\begin{aligned}\frac{dx_{3,i}}{dt} &= -Ax_{3,i} + (B - x_{3,i})(x_{1,i} + G_1 w_{r,i} r + G_3 f(x_{3,i} - \theta)) \\ &\quad -x_{3,i}(x_{2,i} + G_2 w_{p,i} p + H \sum_{j \neq i} f(x_{3,j} - \theta)) \end{aligned} \tag{30}$$

where the excitation terms consist of weighted reward signal multipled by the gain G_1, the excitatory input from the gated dipole and a self feedback term $G_3 f(x_{3,i} - \theta)$. The inhibitory terms consist of weighted punishment signals $G_2 w_{p,i} p$, the inhibitory input from the gated dipole $x_{2,i}$ and the lateral inhibitory term $H \sum_{j \neq i} f(x_{3,j} - \theta)$. The function f is a sigmoid function and is given as

$$f(x) = \frac{x^5}{(1 + x^5)} \tag{31}$$

The differential equations representing the reward and punishment node activities were changed from additive type to the shunting type network. This change causes the effect of the punishment and the reward signals (i.e rise time of the reward/punishment node) to be much faster than the recovery from the signal (i.e the decay time of the reward/punishment node).

Equations representing the activity of the remaining modified gated dipole nodes are given below.

$$\frac{dx_{1,i}}{dt} = -Ax_{1,i} + (B - x_{1,i})(I + J_i) z_{1,i} - x_{1,i} I z_{2,i} \tag{32}$$

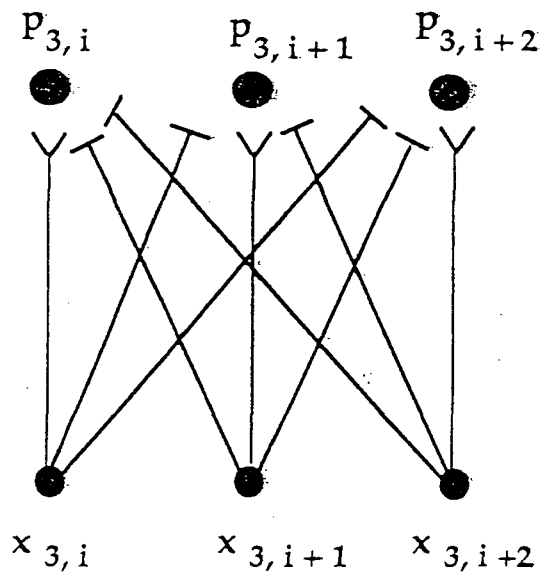
$$\frac{dx_{2,i}}{dt} = -Ax_{2,i} + (B - x_{2,i}) I z_{2,i} - x_{2,i}(I + J_i) z_{1,i} \tag{33}$$

$$\frac{dz_{1,i}}{dt} = \alpha(\beta - z_{1,i}) - \gamma(I + J_i) z_{1,i} \tag{34}$$

$$\frac{dz_{2,i}}{dt} = \alpha(\beta - z_{2,i}) - \gamma I z_{2,i} \tag{35}$$

$$\frac{dw_{r,i}}{dt} = -A_2(w_{r,i} - 1) + (M_r - w_{r,i}) B_2 g(r - \theta_1) g(x_{3,i} - \theta) \tag{36}$$

This concludes the modifications done to the original Levine-Prueitt model.



Synapses notation	Type
—≺	Excitatory
—|	Inhibitory

Figure 14: The Decision Network comprising of a feed-forward shunting type network.

4.5 Simulation Results

4.5.1 Simulations on Gated Dipole

Figure 15a shows a gated dipole neural network. As can be seen from the figure there are two competing channels. A nonspecific arousal input "I" is given to both the channels. When the input "J" to the gated dipole turns on, the greater input to the left of the channel causes x_3 to be excited. Turning off the input J causes a rebound activity to the right channel. The graph showing the input J and output activity of the x_3 neuron are give in Fig 15b.

Each of the gated dipole in the modified model represents an object in the scene. Turning "on" a particular J_i (i.e input to gated dipole) implies that a new object represented by the gated dipole has been introduced in the scene. Figure 16a shows the coupling of two gated dipoles in a lateral inhibitory fashion. In the simulation (of Figure 16b), one object is input (i.e J_1 is turned "on") and after a while a second object is introduced (i.e J_2 is turned on) into the environment. The output graph indicates the activity of the $x_{3,i}$ node of the gated dipoles which represents the choice made by the model. The gated dipole node $x_{3,i}$ having the highest activity implies that, that object was chosen by the model. The input graphs represent when a particular object was introduced in the scene. As can be seen from the graph, initially when the first object was input the activity of the x_3 neuron of the first gated dipole increased causing a drop in the activity of x_3 node of the second gated dipole thus suggesting that the model chose the first object. When the second object was input at time "t_i" the activity of x_3 node of the second gated dipole increased causing the activity of x_3 node of the first gated dipole to drop. The greater activity of the second gated dipole implied that the model had chosen the new object over the other. Thus these simulation demonstrates a competing situation between input objects in which the most recently "input" object was chosen.

4.5.2 Simulations on the new modified model

Figure 17 illustrates three simulations conducted on the new modified model shown in Figure 12. Simulation results shown are for a model comprising of two interacting gated dipoles coupled to reward and punishment nodes. Similar to the above simulation, each of the gated dipoles represents an object in the scene. For each pair of graphs shown, the graph on the left represents the activity of the x_3 node of the first gated dipole and the graph on the right represents the activity of the x_3 node of the second gated dipole. The gated dipole with the highest activity implies, that particular object has been chosen by the model. The first graph illustrates how the model handles novel objects. Initially at $100 timeunits$[8]

[8] The units of time is not crucial however the relative units of time between different time constants are important hence we have refrained from using any specific time units *(e.g secs, msecs etc)* and just referred to it as *"time units"*.

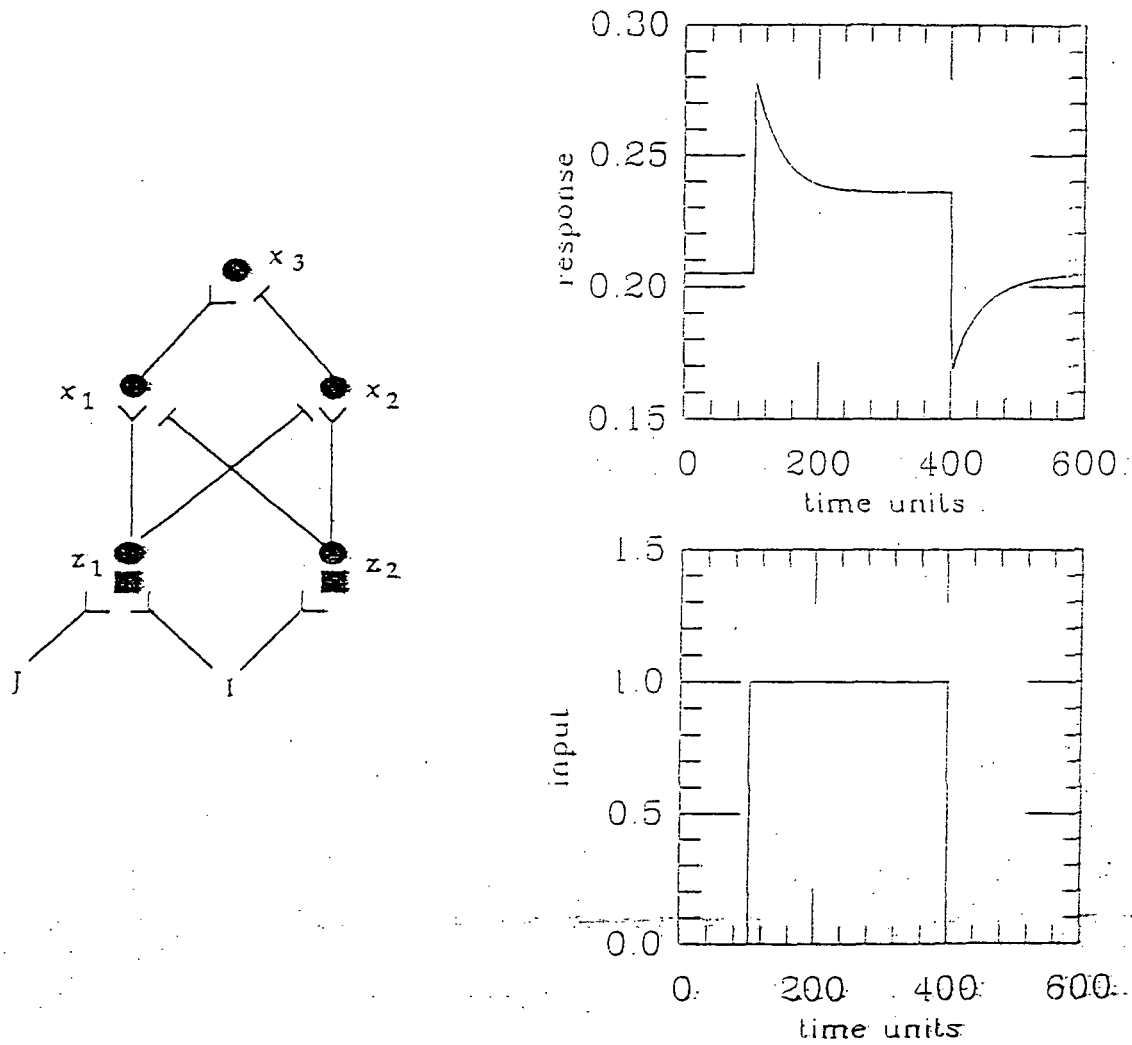
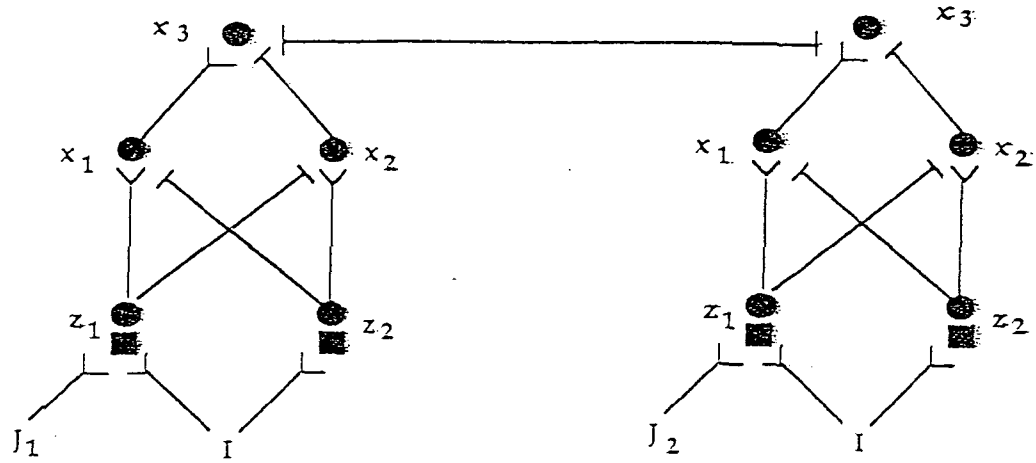


Figure 15: a) Represents a simplified gated-dipole network and b) Shows the dynamics of such a network for an rectangular pulse.



Synapses notation	Type
	Excitatory
	Inhibitory
	Transmitter Gated Synapse
	Represents

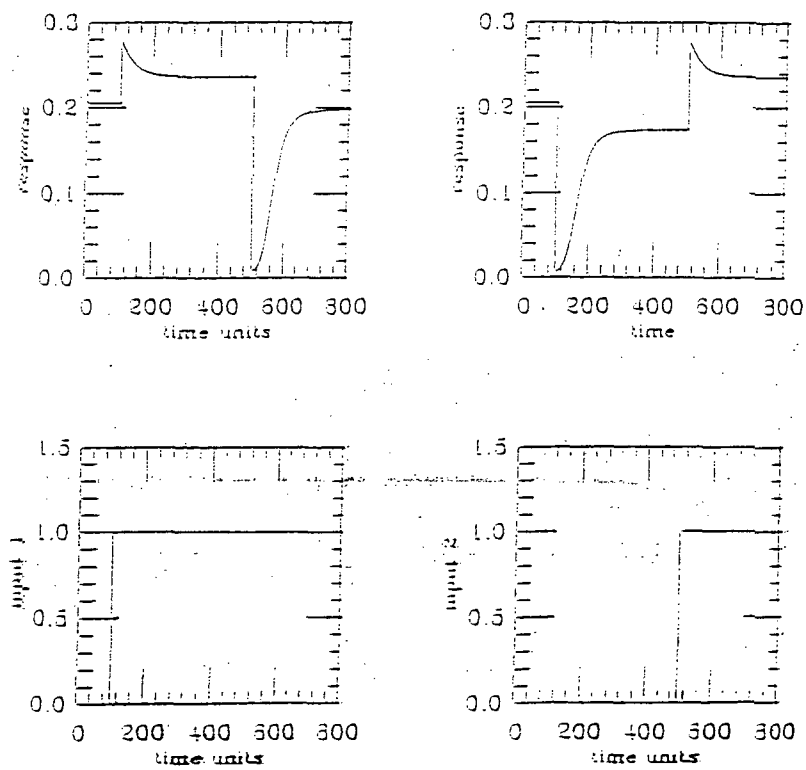


Figure 16: a) Represent a schematic of the coupling between two gated dipoles. b) Shows the dynamics of the system when the inputs J_1 (left) and J_2 (right) are turned on one after another.

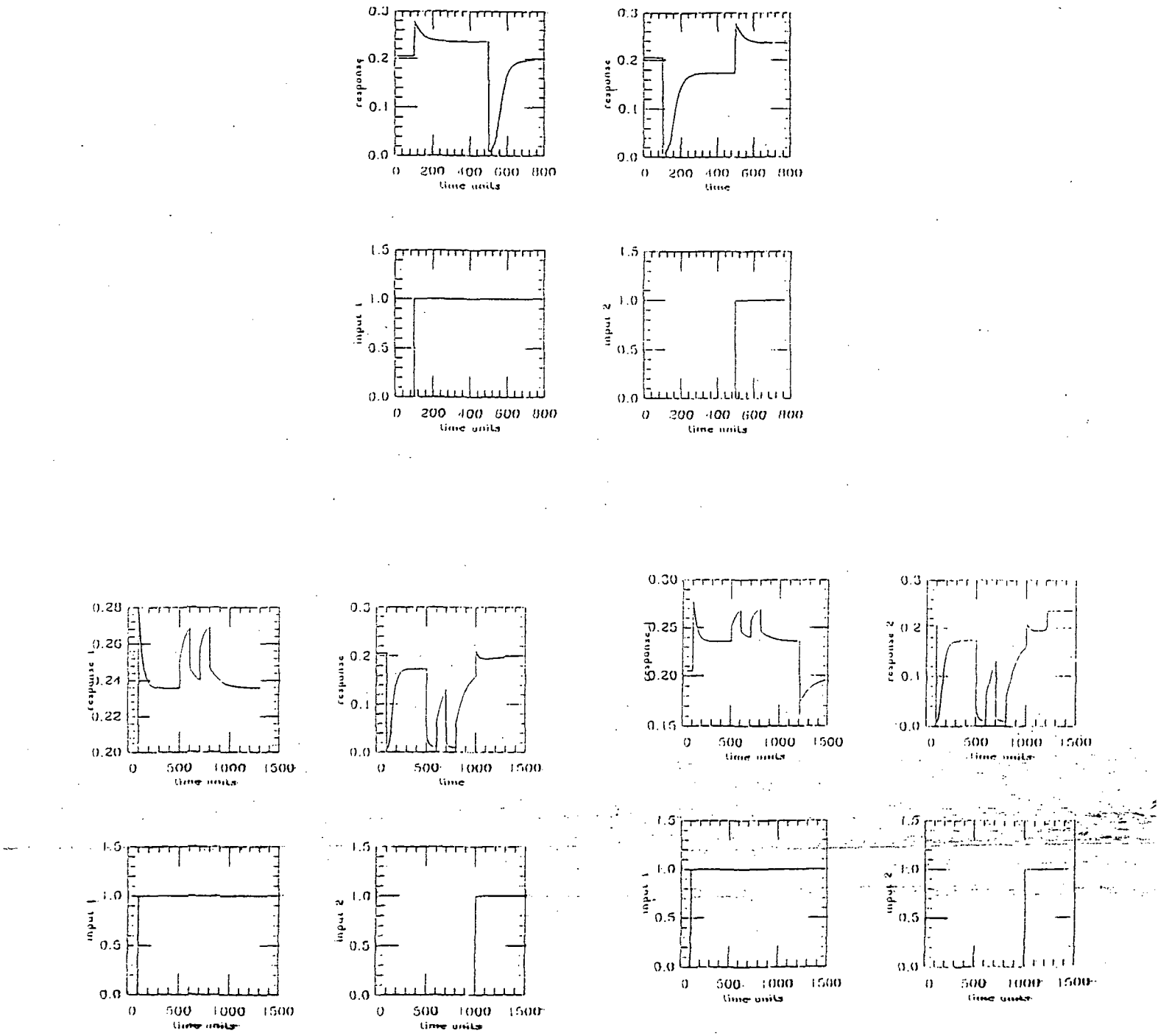


Figure 17: The three graphs show the simulation involving novelty, reward and punishment. Refer text for details.

an object is introduced. This causes the first gated dipole's activity to rise above threshold implying that the model selected this new object. At 800*timeunits* a second object is introduced. By this time the first object has lost its novelty. This leads to the activation of the second gated dipole and which in turn overrides the activity of the first gated dipole. Thus the higher activity of the second gated dipole indicates that the model chose the second object (Note that this occurred when no reinforcement either positive or negative was applied). This simulation shows how most recent input is more attractive than previously unrewarded inputs (**novelty**).

The second pair of graphs demonstrates how previously rewarded objects are more attractive than recently input objects. The simulation begins similar to the previous case i.e an object is input to the scene at 100*timeunits*. However at 500 and 750 *timeunits* the choice of the first object by the model is rewarded. Now when the second object is input at 1000 *timeunits* it fails to drive the second gated dipole above the activity of the first gated dipole implying that the model preferred the first object over the second (novel) object.

The third pair of graphs is a further extension of the previous simulation to include punishment signals. The simulation is similar to the previous case until 1200 *timeunits*.At 1200 *timeunits* punishment signal is given for the choice made by the model. This lowers the activity of the first gated dipole and the second gated dipole response rises above that of the first gated dipole. Thus punishing the initial choice made by the model cause it to change its choice. In the above cases the punishment and reward signals were brief pulses which lasted for 100 *timeunits*.

5 The combined model

We will now describe the combination of these modules with a vision module. For convenience of presentation we will consider a general scenario dealing with the initiation of the exploratory behavior. Such a scenario (Figure 18) is given below. A vision unit is represented by the camera which is looking over a two dimensional terrain; a robotic arm performs certain actions on this terrain; and a cognitive unit coordinates the camera and the robotic arm. The camera concentrates on a particular region and identifies a given object in its region of focus. The cognitive unit categorizes various input objects into a hierarchy of categories, and makes behavioral decisions regarding the object in focus. These behavioral decisions are modulated by novelty, habits and external reinforcement signals. Behavioral decision units interact with object-representation units and a signal is send to the arm as to "reach" for the object whenever appropriate. The schematic view of the various process involved and the interactions between them are given in Figure 19a. Figure 19b gives a more detailed block diagram of this system. All further discussions are related to Figure 19. We begin first from the vision unit.

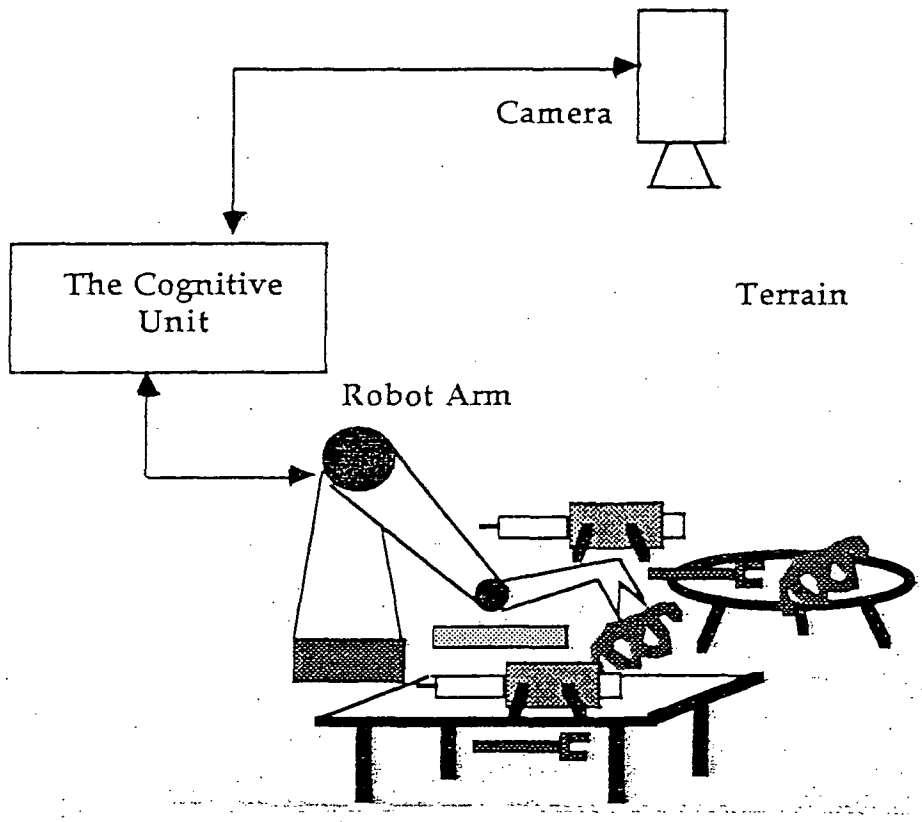
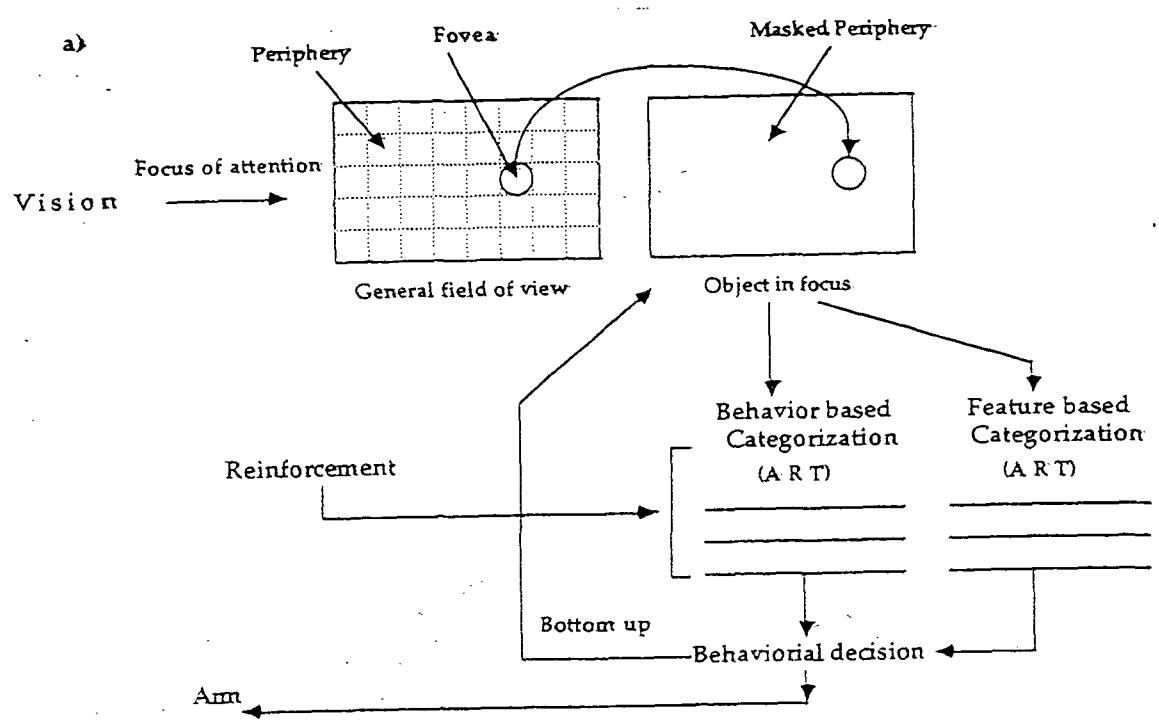


Figure 18: The general scenario for the combined system.



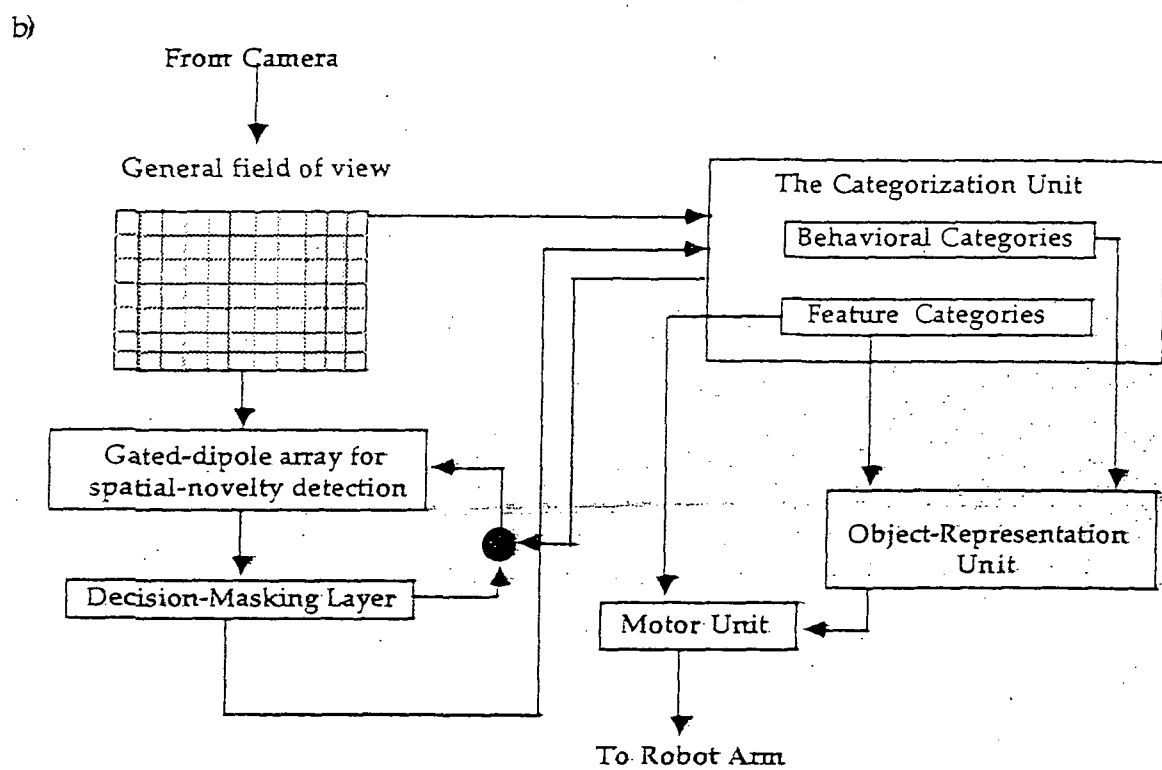


Figure 19: (a) The schematic of the combined system. (b) A more detailed block diagram of the various units in the system.

The general field of view of the camera consists of a topographic array [9]. The neural network for this array is a simplified and modified version of the network proposed in (Öğmen and Gangé 1990a, 1990b; Öğmen and Moussa, inprint). It comprises of a gated-dipole structure that senses temporal novelty of the inputs[10]. This early visual network is shown in Figure 20.

The camera movements are controlled by a competitive layer and a buffering layer as shown in Figure 21. The CL is a recurrent competitive layer. The output of this layer is buffered by a feedforward shunting network (BL) which guarantees that the movements of the eye will not be affected by the transients in the decision layer network (CL). However, such a recurrent competition layer has also the property of hysteresis, i.e the wining neuron will tend to persist and thus will prevent consequent eye movement thereby making the system unable to generate continuous eye movements. To generate continuous eye movements, the wining neuron is inhibited by a recurrent feedback coming from the FL and DL layers. The FL layer combines sensorial and cognitive signals to regulate the eye movement and the DL network adjusts the time course of feedback inhibition to generate a new target position for the eye. Thus this network is responsible for moving the camera to the region of focus. The region of focus is very similar to the fovea in humans, and its the region of high acuity. We have not introduced size and rotation invariance (via e.g. log-polar transform) because we will test the hypothesis that the invariances are learned through mental internalization of sensory-motor behaviors and emerge from exploratory activity.

The foveal input is fed to the categorization networks. ART networks as shown in Figure 22 are chosen to ensure stable categories in non-stationary environments. The ART network at the bottom categorizes the inputs into different object types. The other ART network categorizes the inputs according to categories that are related to the task considered in the present scenario (e.g targeted object, junk object). The categorization of the latter ART network is modulated by bias nodes. The bias nodes, as discussed before, encode the environmental cues which modulate the categorization. The outputs of these categorization networks are fed to another network which combines object novelty and category information to produce a motor behavior. This is shown in Figure 23.

Object types are fed into a gated dipole structure to sense the "object novelty". Slow interneurons are used to integrate transient signals generated by rapid eye movements. The outputs of these gated dipoles fed into a competitive layer with recurrent on-center off-surround connections. This layer also receives outputs from the task related category neurons. The competition generates a winner for the target arm position. Again this competitive layer is buffered by a feed-forward on-center off-surround

[9] This represents the retinotopic organization of the visual space.

[10] We would refer to this novelty as spatial-novelty from here on to differentiate it from object novelty that will be discussed later

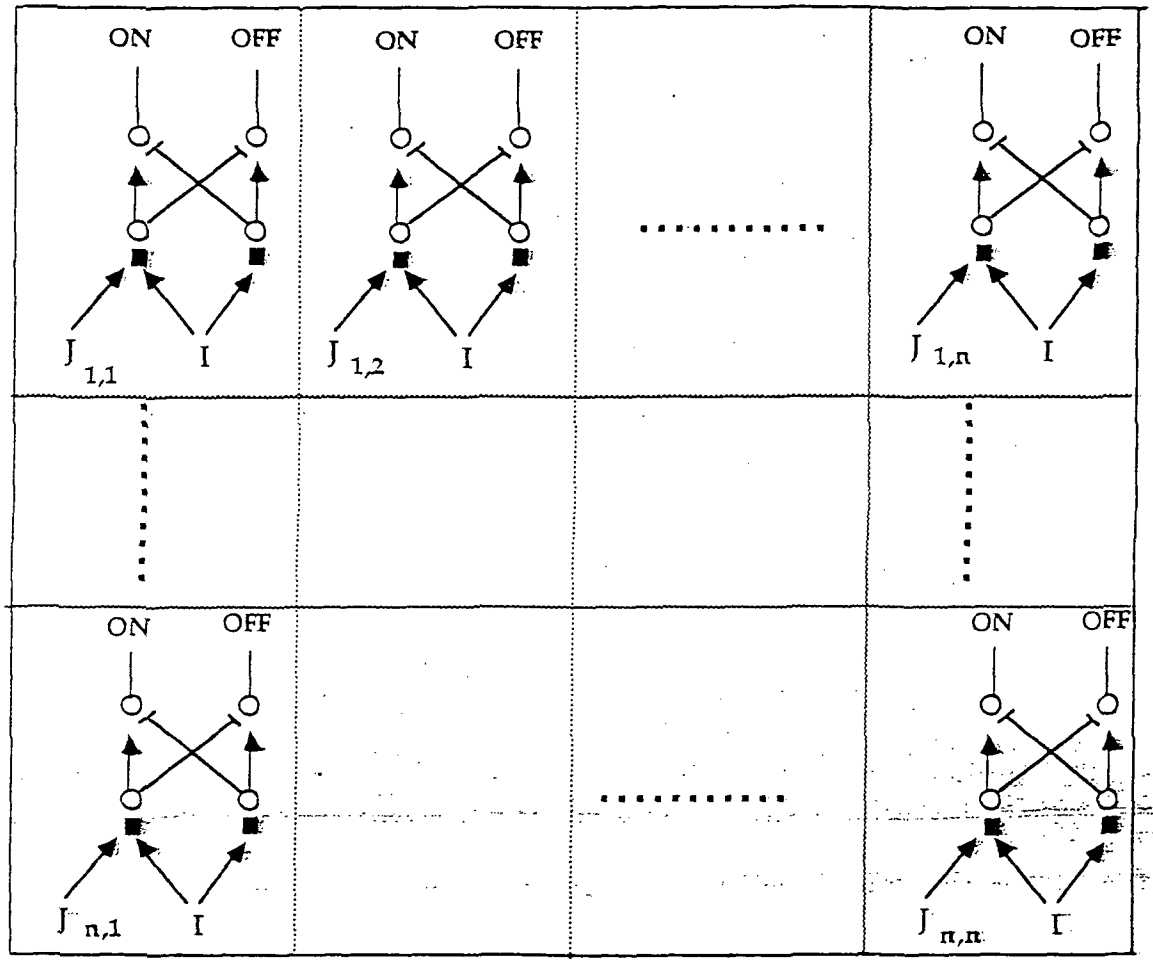


Figure 20: The early visual network modeled by an array of gated dipoles.

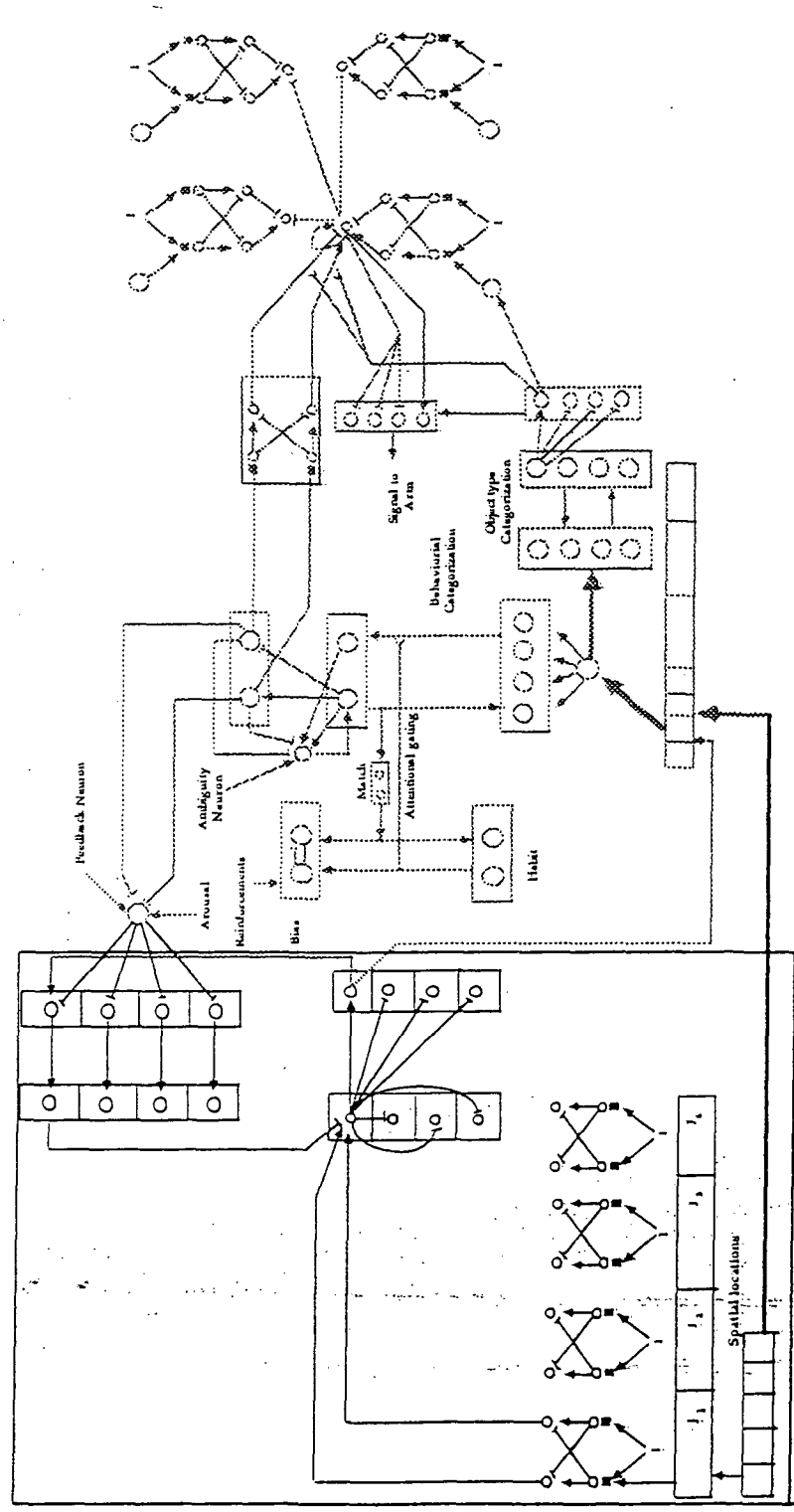


Figure 21: The decision layer that brings the camera "in focus" on a particular object.

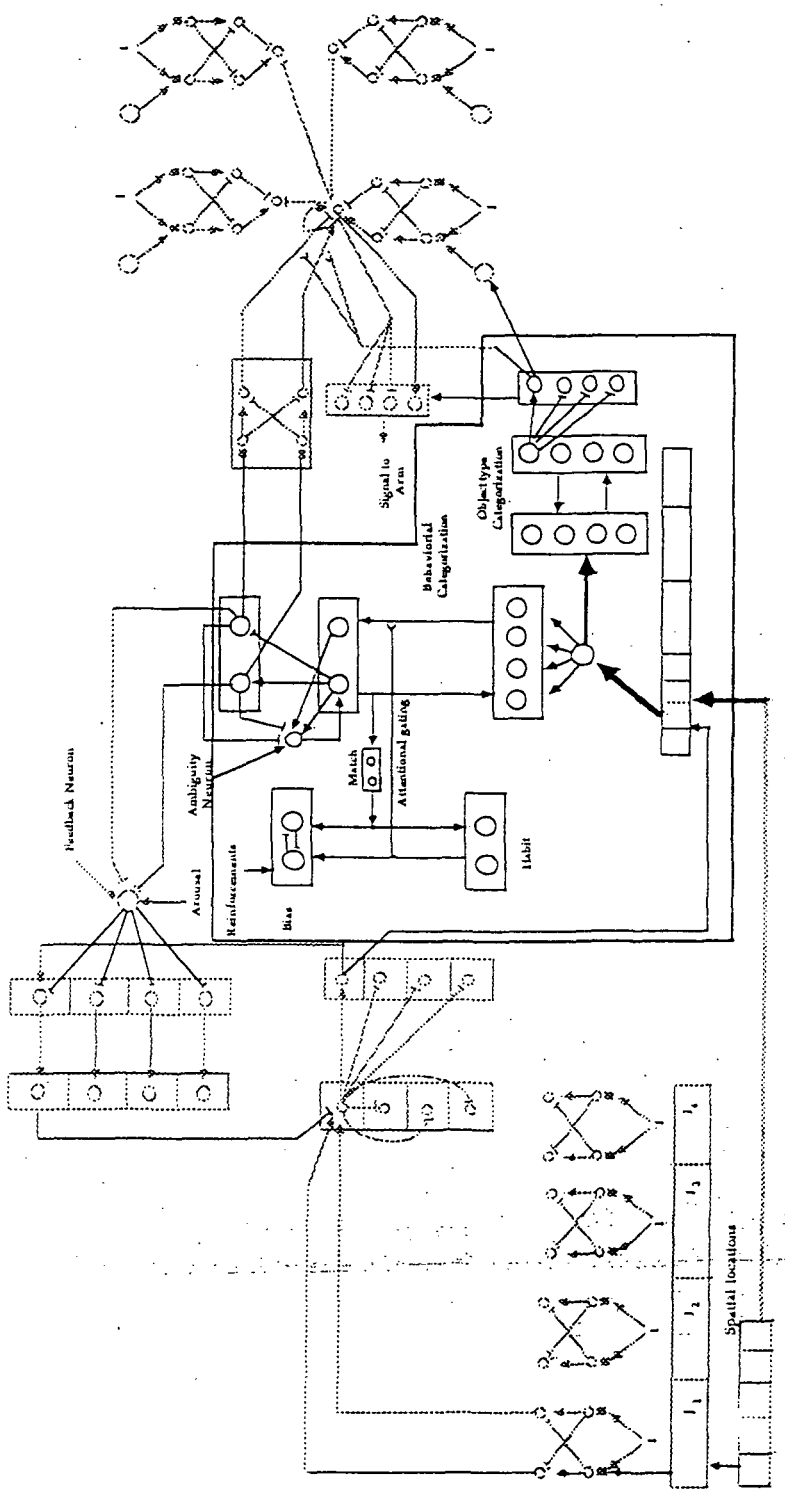


Figure 22: The Categorization unit comprising of two ART networks. The first categorization is based on the features of the objects. The second categorization is based on behavioral significance of the objects.

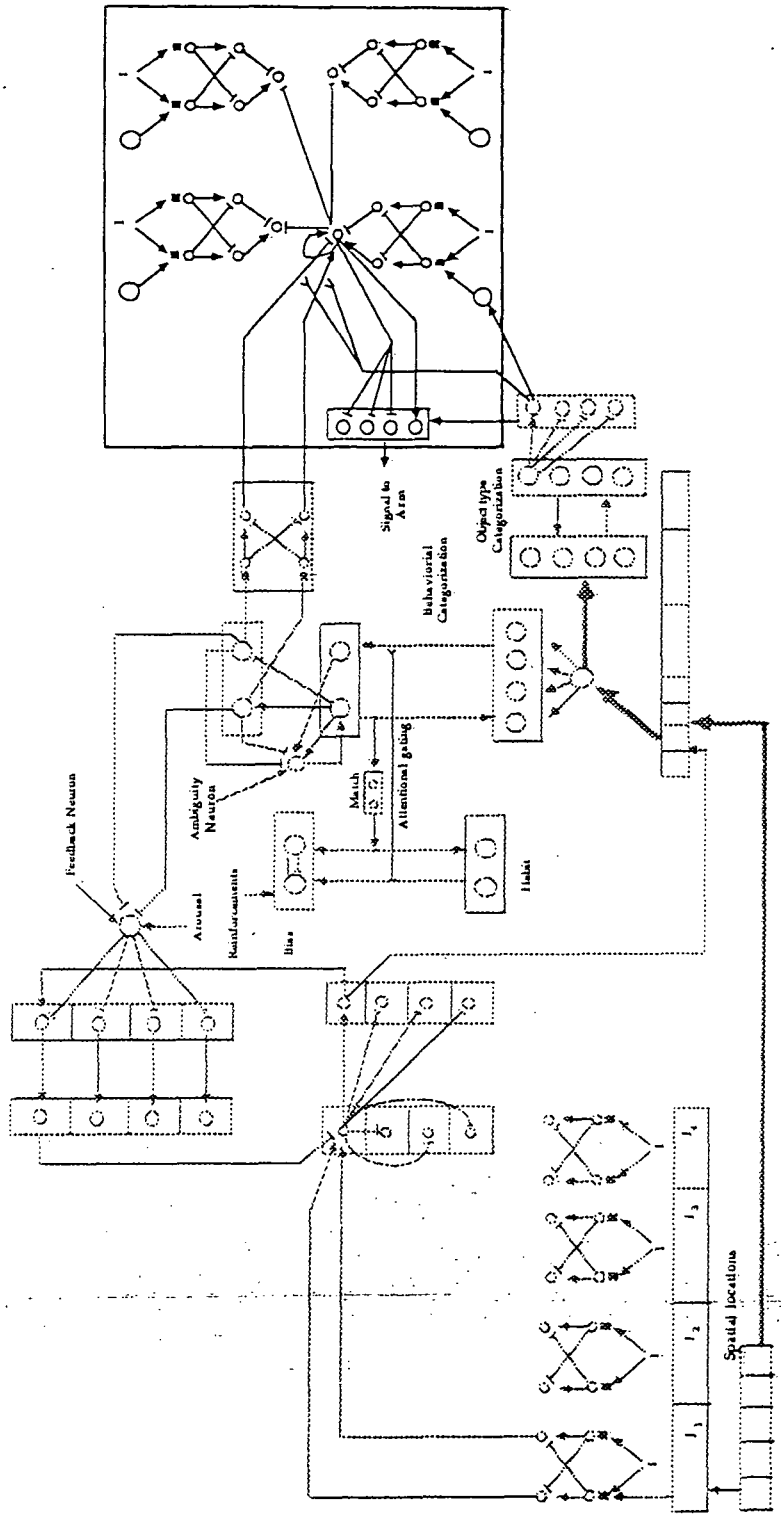


Figure 23: The object-representation unit is shown in the above figure. The inputs "J" to the gated dipole come from the feature categorization layer of the categorization unit. The inputs from the behavioral categorization layer modulate the activity of the recurrent layer neurons in this unit.

network to filter the transients. This buffering layer is gated by an eye position signal to transfer eye position to arm position via a eye-hand coordination network. Thus the combined model would be capable of dynamically locating objects present in its environment and to issue a signal to the robot arm to pick up the object if it is a "targeted object".

5.1 Ambiguity in decision making

In this section we will discuss two combined networks that resolve decisional ambiguities. The first combined architecture is shown in Figure 24. As discussed before, input features of the object in focus are fed to the ART networks. Initially when an object is present in the environment, no behavioral significance is attached to it and an ambiguity exists for the classification of the object as targeted or non-targeted object. The architecture of Figure 25 resolves this ambiguity by biasing the targeted object via the ambiguity neuron. This "positive bias" is used whenever there is insufficient evidence to reach a decision. On the other hand, such a bias forces the system to make a decision for all inputs. A consequence of this, is that novelty plays no role in the decision making, since it is always overridden by the task related decision.

A modification to the above architecture is presented in Figure 26. Note that the inputs that drive the ambiguity neuron in this case come from the decision layer neurons of the object representation network. This modification enables the novelty of the object to play a part in the behavioral decision making process. Now consider the same situation as stated before where an ambiguous object is placed in the environment. In the present situation when a behavioral indecision occurs, the competitive layer feeding to the motor circuits receive their major inputs from the gated dipoles representing object novelty. This signal in turn drives the ambiguity neuron which in turn biases the "targeted" category neuron thus proclaiming the object in focus to be a targeted object. Thus in this modified combined architecture novelty play an important role in cognitive decision making.

In the next subsection we will present the various differential equations that describe the combined model.

5.2 Combined neural model

As mentioned above the combined model comprises of three major networks: the visual scanning network, the cognitive network and the object representation network. The two models discussed above differ only in the implementation of the ambiguity neuron. Hence we will first discuss the differential equations common to both models and finally discuss the ambiguity neuron equation for each case.

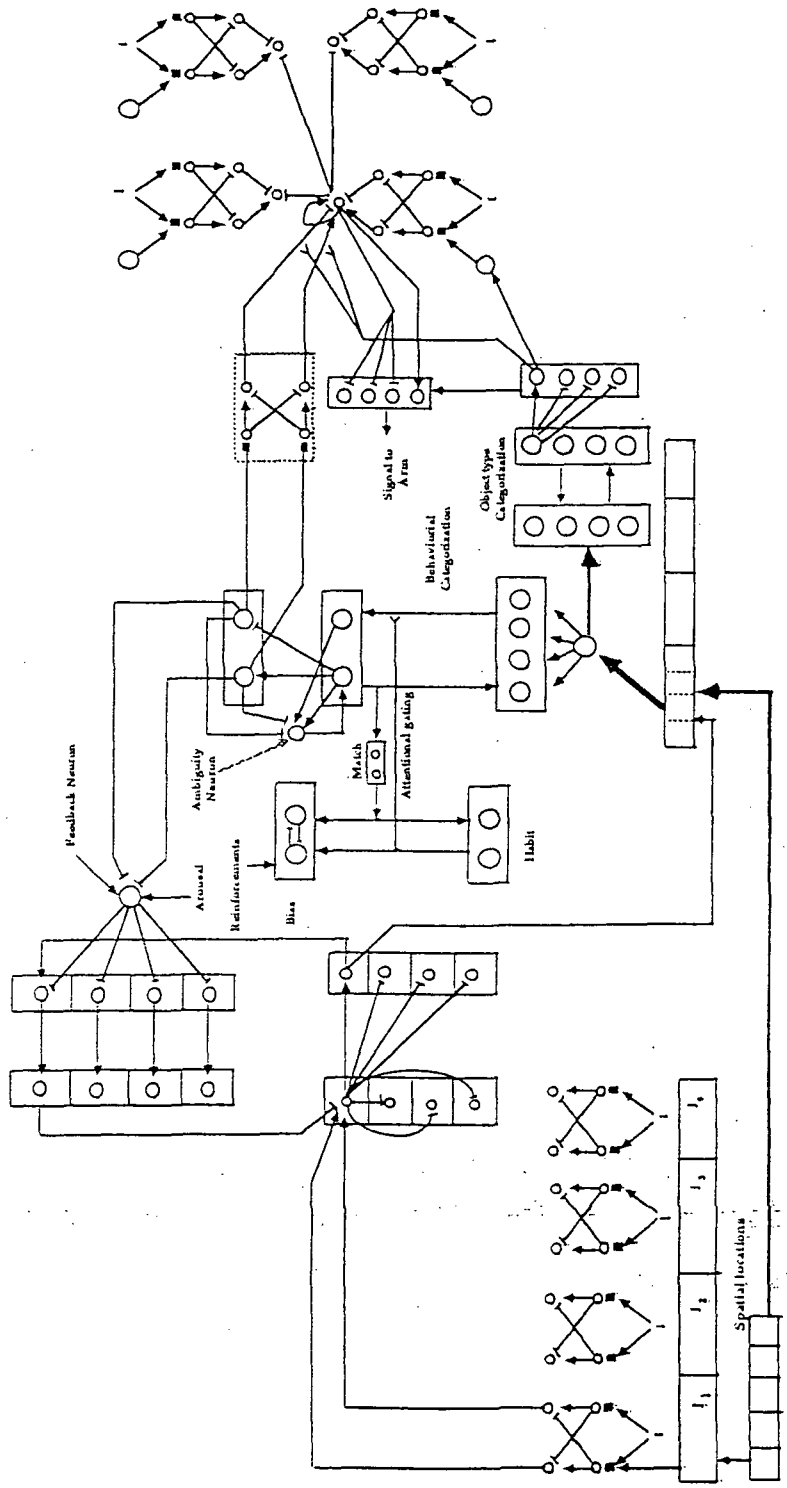


Figure 24: The detailed architecture for the combined model

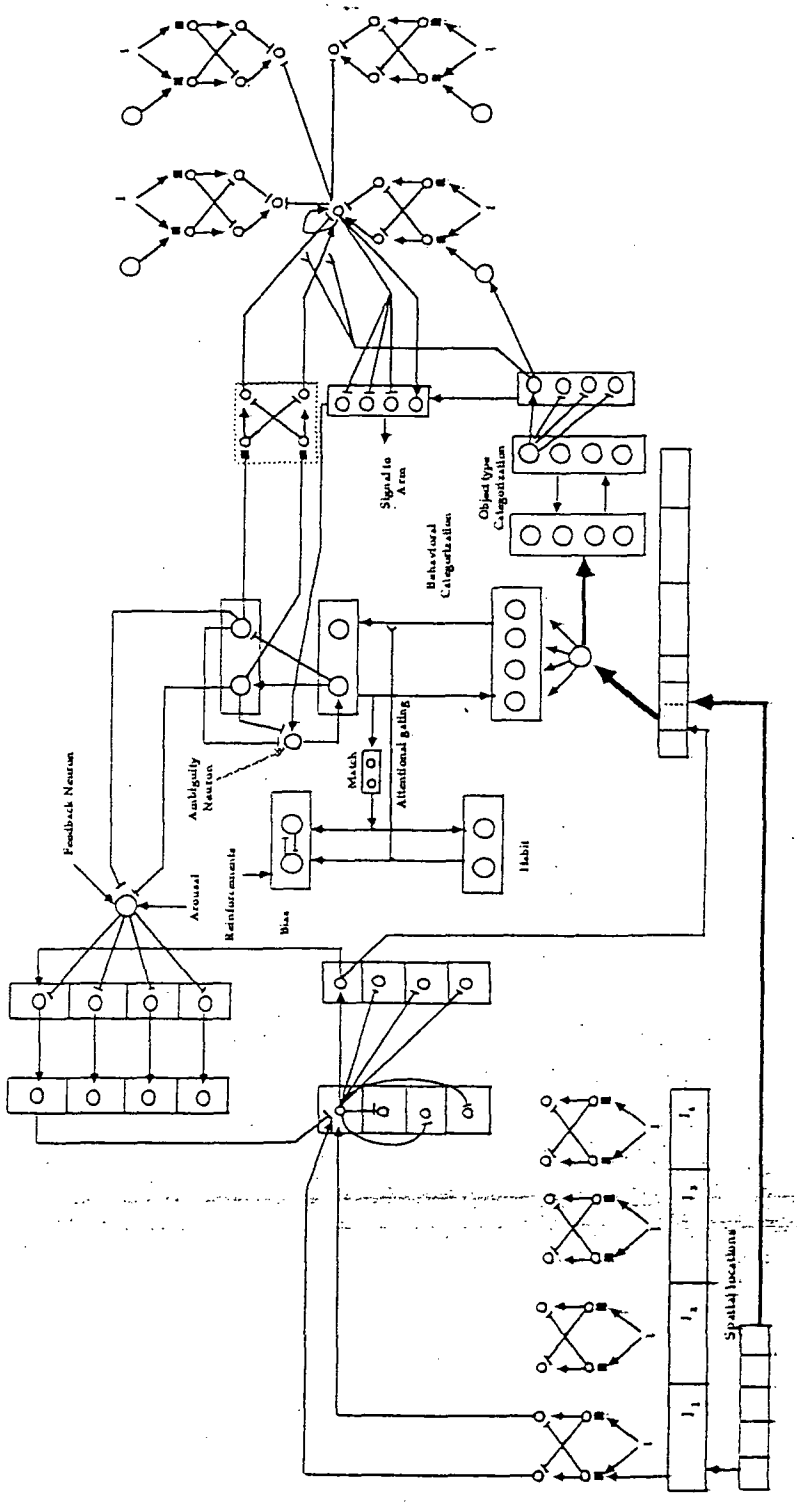


Figure 25: The modified architecture for the combined model.

The visual scanning network is implemented by an array of gated dipoles described by

$$\frac{dvx_{1,i}}{dt} = -Avx_{1,i} + (B - vx_{1,i})(I + J_i)vz_{1,i} - vx_{1,i}Ivz_{2,i} \tag{37}$$

$$\frac{dvx_{2,i}}{dt} = -Avx_{2,i} + (B - vx_{2,i})Ivz_{2,i} - vx_{2,i}(I + J_i)vz_{1,i} \tag{38}$$

$$\frac{dvx_{3,i}}{dt} = -Avx_{3,i} + (B - vx_{3,i})(vx_{1,i} + vx_{2,i} + G_3 f(vx_{3,i} - \theta)) - vx_{3,i}(Ml_i + H\sum_{j\neq i} f(vx_{3,j} - \theta)) \tag{39}$$

$$\frac{dvz_{1,i}}{dt} = \alpha(\beta - vz_{1,i}) - \gamma(I + J_i)vz_{1,i} \tag{40}$$

$$\frac{dvz_{2,i}}{dt} = \alpha(\beta - vz_{2,i}) - \gamma Ivz_{2,i} \tag{41}$$

where $vx_{1,i}$, $vx_{2,i}$, $vx_{3,i}$, $vz_{1,i}$ and $vz_{2,i}$ represent the activities of $x_{1,i}$, $x_{2,i}$, $x_{3,i}$, $z_{1,i}$ and $z_{2,i}$ neurons of the modified gated dipole architecture discussed in section 4.3.

The gated dipole mechanism encodes novelty, and the decision for the new eye position is carried out by the recurrent competitive layer (i.e. the $vx_{3,i}$ neurons). Once attentional gating of the input yields a behavioral action, an inhibiting signal is issued by the feedback circuit. This feedback inhibition is represented by l_i neurons. The competitive layer is buffered using a feedforward network so that transient fluctuations during competition are avoided. This buffering layer known as the decision layer provides a signal when a decision has been achieved by the recurrent layer. The output of the this decision layer is used to control the eye movement. The decision layer performs the function of moving the eye to a particular object and masking peripheral view while categorizing the objects. The differential equation for this feedforward shunting decision layer is given below.

$$\frac{dp_{1,i}}{dt} = -Ap_{1,i} + (B - p_{1,i})Wvx_{3,i} - p_{1,i}W\sum_{j\neq i} vx_{3,j} \qquad i,j = 1,2,..15. \tag{42}$$

The output of this layer is thresholded. The value of the threshold reflects the certainity in the decision (e.g with B = 1, 0.9 corresponds to 90% of votes). The behavioral and the object categorization ART networks categorize the attentionally gated visual input and the equation for the respective networks are given below. The differential equations for the behavioral ART is given as follows

$$\frac{dbx_i}{dt} = -Abx_i + (B - Cbx_i)(I_i + \sum_{j=1}^{4} f(by_j)bz_{j,i})$$
$$-Dbx_i\sum_{j=1}^{4} f(by_j), \qquad i = 1,2,3,4. \tag{43}$$

$$\frac{dby_j}{dt} = -Aby_j + (B - Cby_j)(f(by_j) + \sum_{i=1}^{12} g(\Omega_{[\frac{i+3}{4}]}bx_i)bw_{i,j})$$

$$-Dby_j(\sum_{r\neq j} f(by_r) + \mathcal{I}), \qquad j = 1, 2. \tag{44}$$

And the differential equations for the object categorization network are as follows.

$$\begin{aligned}\frac{dfx_i}{dt} &= -Afx_i + (B - Cfx_i)(I_i + \sum_{j=1}^{4} f(fy_j) fz_{j,i}) \\ &\quad -Dx_i \sum_{j=1}^{4} f(fy_j), \qquad i = 1, 2, 3, 4. \end{aligned} \tag{45}$$

$$\begin{aligned}\frac{dfy_j}{dt} &= -Afy_j + (B - Cfy_j)(f(fy_j) + \sum_{i=1}^{12} g(fx_i) fw_{i,j}) \\ &\quad -Dfy_j(\sum_{r\neq j} f(fy_r) + \mathcal{I}), \qquad j = 1, 2, 3, 4. \end{aligned} \tag{46}$$

where bx_i, by_j, fx_i and fy_j are the feature and category neuron activities for the behavioral and object categorization networks respectively.

The transients that arise during competition in the F2 layer of the object i and behavioral categorization network are also buffered using a decision layer. The differential equation for the feedforward shunting decision layer is as follows.

$$\frac{dp_{2,i}}{dt} = -Ap_{2,i} + (B - p_{2,i}) W by_i - p_{2,i} W \sum_{j\neq i} by_j \qquad i, j = 1, 2. \tag{47}$$

$$\frac{dp_{3,i}}{dt} = -Ap_{3,i} + (B - p_{3,i}) W fy_i - p_{3,i} W \sum_{j\neq i} fy_j \qquad i, j = 1, 2, 3, 4. \tag{48}$$

Internal biases (due to habit) and external biases (due to reinforcement) that influence the categorization of the behavioral ART network are incorporated in the bias nodes of the cognitive system. The differential equations for the bias and the habit nodes are given below.

$$\begin{aligned}\frac{d\Omega_k}{dt} &= -E\Omega_k + \{(F - \Omega_k)([h_k - \theta_1]^+ + \alpha R^+ + g(\Omega_k)) \\ &\quad -\Omega_k(\alpha R^- + G \sum_{r\neq k} g(\Omega_r))\} f(\Phi_k) \qquad k = 1, 2. \end{aligned} \tag{49}$$

$$\frac{dh_k}{dt} = Hh_k\{(J - h_k)[\Phi_k - \theta_2]^+ - [\Phi_k - \theta_2]^-\} \qquad k = 1, 2. \tag{50}$$

where Ω_k and h_k represent the activities of the bias and habit neurons respectively. R^+ and R^- are the positive and negative environmental cues and Φ_k is the match signal, similar to that discussed in section 3.3.

Once a relevant behavioral categorization has been achieved the feedback neuron is inhibited which also signals the eye to initiate scanning. Thus scanning is achieved by inhibiting the currently active neuron so that other neurons have the possibility to win the competition. The eye then moves,.to the object present at that location. The scanning is achieved by the following two layer of neurons and the feedback neuron. The differential equations for this system are as follows.

$$\frac{df}{dt} = -A_f f + (B - f)Arousal - fH\sum_j g_1 p_{2,j} \qquad j = 1, 2. \tag{51}$$

$$\frac{ds_i}{dt} = -A_s s_i + (B - s_i)G_1 g(p_{1,i} - \theta_1) - s_i H_1 g(f - \theta_2) \qquad i = 1, 2, ..15. \tag{52}$$

$$\frac{dl_i}{dt} = -A_l l_i + (B - l_i)G_2 g(s_i - \theta_2) \qquad l = 1, 2, ..15. \tag{53}$$

where f represents the activity of the feedback neuron, s_i represents the activity of neurons in the feedback layer and l_i represents the activity of neurons in the delay layer, which provides appropriate synchrony off the feedback signals.

Behavioral decision influences the object representation network which in turn signals the robot arm whether or not to pick the currently focused object. When no definite behavioral categorization is achieved, novelty of object classes plays an important role. The object representation system too comprises of a gated dipole network so as to model the novelty associated with the various classes of objects present in the environment. The differential equations for the object representation system are given below.

$$\frac{dcx_{1,i}}{dt} = -Acx_{1,i} + (B - cx_{1,i})(I + q_i)cz_{1,i} - cx_{1,i}Icz_{2,i} \tag{54}$$

$$\frac{dcx_{2,i}}{dt} = -Acx_{2,i} + (B - cx_{2,i})Icz_{2,i} - cx_{2,i}(I + q_i)cz_{1,i} \tag{55}$$

$$\begin{aligned}\frac{dcx_{3,i}}{dt} = & -Acx_{3,i} + (B - cx_{3,i})(cx_{1,i} + G_3 f(cx_{3,i} - \theta) + G_3 p_{2,1}) \\ & -cx_{3,i}(cx_{2,i} + H\sum_{j\neq i} f(cx_{3,j} - \theta) + G_3 cp_{2,2})\end{aligned} \tag{56}$$

$$\frac{dcz_{1,i}}{dt} = \alpha(\beta - cz_{1,i}) - \gamma(I + J_i)cz_{1,i} \tag{57}$$

$$\frac{dcz_{2,i}}{dt} = \alpha(\beta - cz_{2,i}) - \gamma Icz_{2,i} \tag{58}$$

$$\frac{dq_i}{dt} = -A_q q_i + (B - q_i)p_{3,i} \tag{59}$$

where, similar to the previous case $cx_{1,i}$, $cx_{2,i}$, $cx_{3,i}$, $cz_{1,i}$ and $cz_{2,i}$ represent the activities of various neurons of the modified gated dipole network. q_i represents the activity of the slow neuron which

encodes the various object types present in the environment. This neuron is also involved in buffering the transients produced due to the rapid scanning of the field of view by the robot.

The recurrent layer in the object representation system also gives rise to transients which need to be buffered before they are sent to generate arm signals. If this is not done, the arm precariously moves around during the decision process. Hence the output of the buffered decision layer is used to generate the arm signal. The differential equation for this shunting feedforward type decision layer is as follows.

$$\frac{dp_{4,i}}{dt} = -Ap_{4,i} + (B - p_{4,i})Wcx_{3,i} - p_{4,i}W\sum_{j\neq i} cx_{3,j} \qquad i,j = 1,2,..4. \tag{60}$$

Sometimes a behavioral decision cannot be made regarding the object currently in focus. This can be due to conflicting environmental cues or may be due insufficient knowledge. As discussed in the previous section this ambiguity can be resolved in two different ways depending on the behavioral significance.

$$\frac{da}{dt} = -Aa + (B - a)\sum_j bx_j - a\sum_j p_{2,j} \qquad j = 1,2. \tag{61}$$

$$\frac{da}{dt} = -Aa + (B - a)\sum_k p_{4,j} - a\sum_j p_{2,j}$$
$$j = 1,2. \qquad k = 1,2,3,4. \tag{62}$$

where $p_{2,j}$ and $p_{4,j}$ are the decision layer neurons of the behavioral categorization and object representation networks.

The first of the above two equations represents the behavioral scenario where an object in focus is deemed as good if no behavioral categorization could be achieved, thus initiating the arm to go and pick the object. The second equation represents the scenario in which novelty of the object initiates a "pick" signal to the arm when a decisive behavioral categorization is not possible.

5.3 Simulation Results

A simulation showing the spatial scanning performed by the robot on two object in its field of view is shown in Figure 26. The curves in the Figure graphs the activity of the visual neuron $vx_{3,i}$ (which represents the robots focus of attention). The continuous and the dashed curves thus represent the activities due to the two respective objects. Similarly the two different kind of spikes illustrate whether the object in focus has been categorized as a targeted object (represented by a continuous spike) or as a junk object (represented by dashed spike).

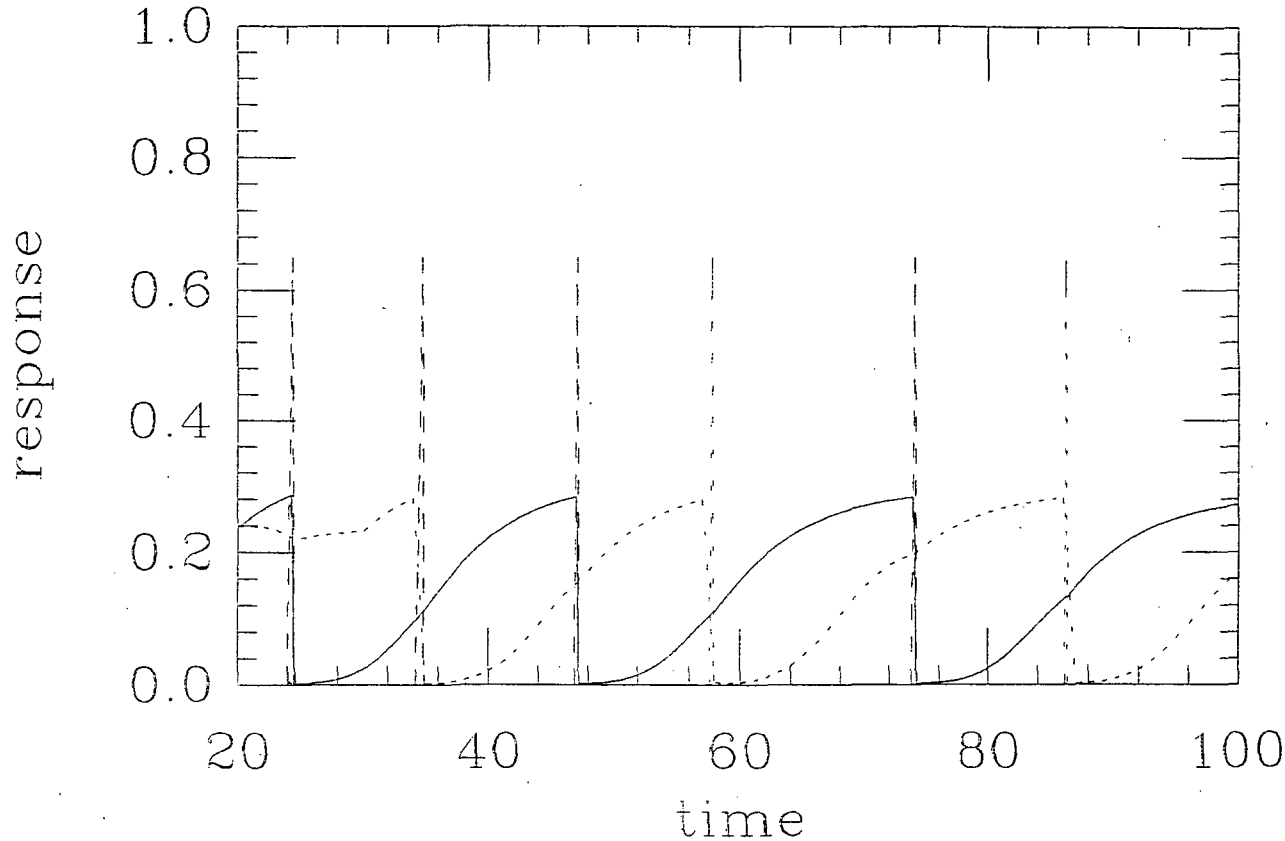


Figure 26: Simulation for the scanning and categorization of two objects are shown in the above graph. The feature vectors for these two objects are (1 0) and (0 1) respectively. The two kinds of curved graphs: the continuous and the dashed represent the object that is currently in focus. The two types of spikes represent the categorization of the object into the target object category (the dashed line) and the junk category (sparse dashed lines). Initially when first object is introduced into the environment. The robot focuses on the object and categorizes it as a targeted object. A positive reinforcement is given to suggest that the decision of the robot was indeed correct. Subsequently the robot focuses on the second object and deems it as a targeted object. A negative reinforcement signal is applied to suggest that the object is a junk object. The robot then focus back and forth between the two objects categorizing them correctly into target and junk objects. Note that three reinforcement signals (yes, no, yes) are required for the robot to correctly categorizes according to first feature. Moreover reinforcement signals are applied only when the an object is considered as a targeted object, and the robot picks it up.

The scenario consists of two objects introduced one after another in the field of view of the robot. Both the objects have complementary types of features i.e one of the objects is a grey wrench and the other is a pink spring. The feature vectors describing the two objects are (1001) and (0110) respectively [11]. Initially a grey wrench is introduced, the robot zeros in on it and categorizes it as a good object. The robot is rewarded for its action. Consequently the spring appears in the robot field of view causing it to focus on this new object. Since there is insufficient evidence regarding the behavioral quality of the object, the robot picks this new object up. Subsequently the robot is punished for its decision [12]. The scanning mechanism then focuses on to the first object and categorizes it as a targeted object. It then scans back to the pink spring but does not pick it up because it too is categorized correctly as an junk object. Thus after three reinforcements the robot learnt to categorize object target object or a junk object depending on whether it is tool or not.

6 Conclusion

We described in this report a neural network based robotic system. Unlike traditional robotic systems, our approach focussed on non-stationary problems. We indicated that self-organization capability is necessary for any system to operate successfully in a non-stationary environment. We suggested that self-organization should be based on an active exploration process. We investigated neural architectures having novelty sensitivity, selective attention, reinforcement learning, habit formation, flexible criteria categorization properties and analyzed the resulting behavior (consisting of an intelligent initiation of exploration) by computer simulations.

While various computer vision researchers acknowledged recently the importance of active processes(Swain and Stricker, 1991) the proposed approaches within the new framework still suffer from a lack of self-organization(Aloimonos and Bandyopadhyay, 1987; Bajcsy, 1988).

A self-organizing, neural network based robot (MAVIN) has been recently proposed (Baloch and Waxman, 1991). This robot has the capability of position, size, rotation invariant pattern categorization, recognition and pavlovian conditioning. Our robot does not have initially invariant processing properties. The reason for this is the emphasis we put on active exploration. We maintain the point of view that such invariant properties emerge from an internalization of exploratory sensory-motor activity. Rather than coding the equilibria of such mental capabilities we are seeking to capture its dynamics to understand on the one hand how the emergence of such invariances is possible and on the other hand

[11] $(\vec{f}_1 \vec{f}_2)$ represents the features of the input object where $\vec{f}_1$ implies whether the object is a tool (10) or not (01) and $\vec{f}_2$ states whether it is pink (10) or grey (01) in color.

[12] Note that the robot is rewarded or punished only when it picks an object.

the dynamics that lead to these invariances. The second point is crucial for an adaptive robot to acquire new invariances in non-stationary environments as demonstrated by the inverting glass experiments of Helmholtz [13].

We will introduce Pavlovian conditioning circuits in our future work for the precise objective of achieving the generation, coordination and internalization of *sequence* of actions. From this perspective, our system differs from MAVIN in that MAVIN passively associates stimuli (as the passive cat of Held's experiment (Held, 1963, 1965)) while our system, in addition to passive association, acquires active association (as the active cat in Held's experiment(Held, 1963, 1965)).

In our future work we will implement this system in hardware. Such an implementation will lead to long term tests of the system. The natural theoretical extension of the work is the integration of associative learning circuits for the generation of *sequences* of actions and their intelligent coordination, and mental internalization.

[13] A similar phenomenon is experienced by many of us who try glasses for the first time. Initially we, are exposed to a distorted perception which is consequently corrected by the brain

References

[1] Aloimonos J., and Badyopadhyay A., Active Vision, *IEEE 1st International Conference on Computer Vision,* pp. 35-54, June 1987.

[2] Bajcsy R., Active perception. *Proceedings of the IEEE,* vol. 76, pp. 996-1004, August 1988.

[3] Baloch A.A., and Waxman A.M., Visual Learning, Adaptive Expectations, and Behavioral Conditioning of the Mobile Robot MAVIN, *Neural Networks,* vol. 4, pp. 271-302, 1991.

[4] Brooks R.A., A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation,* vol RA-2, pp. 14-23, March 1986.

[5] Brooks R.A., A Robot that Walks; Emergent Behaviors from a carefully Evolved Network. *Neural Computation,* vol. 1, pp. 253-262, 1989.

[6] Carpenter G.A., and Grossberg S., A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Processing,* vol. 37, pp. 54-115 1987.

[7] Carpenter G.A., and Grossberg S., The Art of Adaptive Pattern Recognition by a Self-Organizing Neural Network. *IEEE Computer,* pp. 77-88, March 1988.

[8] Elkind D., and Flavell J. H., *Studies in Cognitive Development: Essays in Honor of Jean Piaget.* Oxford University Press, 1969.

[9] Grossberg S., A Neural Theory of Punishment and Avoidance, II: Quantitative Theory, *Mathematical Bioscience,* vol. 15, pp. 253-285, 1972.

[10] Grossberg S., Contour enhancement, short term memory and constancies in reverberating neural networks, *Studies in Applied Mathematics,* vol. 52, pp. 217-257, 1973.

[11] Grossberg S., Nonlinear Neural Networks: Principles, Mechanisms, and Architectures, *Neural Networks,* vol. 1. pp. 17-61, 1988.

[12] Held R., and Bossom, J., Neonatal deprivation and adult rearrangement: Complementary techniques for analyzing plastic sensory-motor co-ordinations. *Journal of Comparative and Physiological Psychology,* vol. 54, pp. 33-37, 1961.

[13] Held R., and Hein, A., Movement-produced simulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology,* vol. 56, pp. 872-876, 1963.

[14] Held R., Plasticity in sensory-motor systems. *Scientific American,* vol. 213, pp. 84-90, 1965.

[15] Leven, S. J., and Levine D. S., Effects of reinforcement on knowledge retrieval and evaluation. *Proceedings of the First International Conference on Neural Networks, San Diego: IEEE/ICNN,* vol. II, pp. 269-279, 1987.

[16] Levine D. S., and Prueitt P. S., Modeling some effects of frontal lobe damage-novelty and perseveration. *Neural Networks,* vol. 2, pp. 103-116, 1989.

[17] Milner B., Effects of different brain lesions in card sorting . *Arch. Neurol,* vol. 9, pp. 90-100, 1963.

[18] Milner B., Some effects of frontal lobectomy in man. In: J.Warren and K.Akert (Eds.), *The Frontal Granular Cortex and Behavior,* New York: MsGraw-Hill, 1964.

[19] Öğmen, H., and Gagné, S., Neural network models for sustained and ON-OFF units of insect lamina. *Biological Cybernetics,* vol. 63, pp. 51-60, 1990a.

[20] Öğmen, H., and Gagné, S., Neural network architecture for motion detection in the fly visual system. *Neural Networks,* vol. 3, pp. 487-505, 1990b.

[21] Öğmen, H., and Moussa, M., A neural architecture for visually guided motor behavior: The landing reaction in flies (in preparation).

[22] Piaget J., *The origins of intelligence in children.* Norton: New York, 1963.

[23] Piaget J., *La psychologie de l'intelligence.* Armand Colin: Paris, 1967.

[24] Piaget J., *The mechanisms of perception.* Basic Books: New York, 1969.

[25] Piaget J., *Psychologie et épistémologie.* Denoël Gonthier: Paris, 1970.

[26] Pribram, K. H., A further experimental analysis of the behavioral deficit that follows injury to the primate frontal cortex. *Experimental Neurology,* vol. 3, pp. 432-466, 1961.

[27] Swain M.J., and Stricker M., (Eds) *Promising Directions in Active Vision* NSF Active Vision Workshop, University of Chicago Technical Report CS 91-27, November, 1991.