A Model for Hormonal Modulation of Learning

Hiroaki Kitano
Sony Computer Science Laboratory
3-14-13 Higashi-Gotanda, Shinagawa
Tokyo 141 Japan
kitano@csl.sony.co.jp

Abstract

Recent studies in neurobiology have discovered that many hormones exist in the brain, and play key roles in learning and memorization. In this paper, we discuss the possible role of hormones in learning, and propose a new learning model which incorporates hormonal effects on learning. The model is a variant of reinforcement learning with modulation on learning rate and the frequency of mental rehearsal. The modulation enables the system to focus its learning on data which are evaluated as important for the system's overall performance. The experiment demonstrates that the incorporation of hormonal modulation improves behavior learning performance, and such an evaluation network can be acquired through the evolutionary mechanism.

1. Introduction

Recent developments in neuroscience and molecular biology have identified several neural systems which discharge hormones and neuropeptides according to emotional and other internal body status. Traditionally, hormones have been thought to control the homeostasis of internal organs, but recent findings suggest that the central nervous system itself produce hormones, thereby affecting behavior and learning. In this paper, we discuss the possible role of hormones in learning, and propose a variant of reinforcement learning. The central claim in this paper is that hormonal modulation of learning (HML) enables the system, or the agent, to focus learning on important incidents for performance and survival, thus improving the total behavioral performance over learning agents without hormonal modulations. Evolutionary acquisition of learning focus is essential for autonomous systems in an open dissipative environment. In such an environment, a priori acquisition of training data and a behavior evaluation function is not possible. In addition, importance of the event is not related with the frequency the event takes place. This paper discusses the use of hormoal modulation of learning for designing autonomous agents in the open dissipative environment.

2. Molecular Neurobiology

This section presents a brief overview of molecular neurobiology relevant to the discussions in this paper. The central issue is how signals among neurons are transmitted and modulated. In the traditional literature, neurotransmitters such as acetylcholine, glutamate, and β-aminobutyric acid are considered to be the sole chemical components mediating inter-neuron information processing. However, recent studies, particularly in the last decade, revealed that many chemical substances known as hormones are playing important roles in the central nervous system. For details, refer to [Hall, 1992; McGaugh, 1989].

2.1. Neurotransmitters

It is well understood that chemical substances mediate signal transmission among neurons. When electric impulses, which propagate through the axon, reach a pre-synaptic site (or axon terminal), chemical substances such as acetylcholine (ACh), glutamate, and γ-aminobutyric acid (GABA) are released into the synapse. ACh and glutamate mediate excitory connections, and GABA mediate inhibitory connections. These neurotransmitters act on ion channels on the surface of the post-synaptic site (Fig. 1-left). The ion channels then open to intake Na^+ which causes elevation of the action potential. This process causes fast excitatory postsynaptic potentials (fEPSF). This part of the information processing has been modeled in current neural networks. However, there are other chemicals involved in the information processing.

2.2. Biogenic amines

The second group of chemicals, the biogenic amines, include serotonin (5-hydroxytryptamine), histamine, norepinephrine, epinephrine, and dopamine. These chemicals are known as hormones, and are used by endocrine and other cells as well as by neurons. Synaptic communication mediated by amines has aspects of both fast and slow transmission. Typically, amines are released by paracrine discharge, which means it diffuses to nearby synapses, as opposed to specialized discharge of neuro-

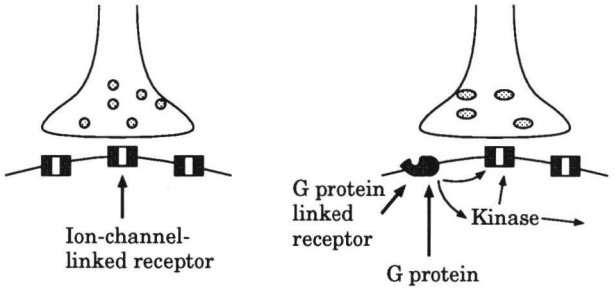

Figure 1: Synapse

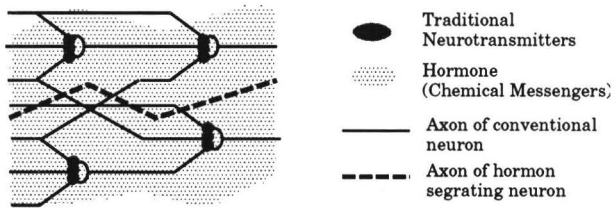

Figure 2: Diffusion of Hormone from an Axon

transmitters which acts only at the synapse where the chemicals are discharged. This is illustrated in Fig 2.

Amines act on G protein-linked receptors, and not directly on the ion channels. The process involves the intra-cellular second messengers such as cAMP and PKC to open the ion channels (Fig 1-right). The signal transmission through this process is slow in both onset (tens of milliseconds) and in duration (possibly a few hours).

Amines are discharged from specialized neural systems called the A, B, and C neural systems. Fig. 3 shows an example of projection from the A_{10} neural system. Many hormone producing neurons exist in the hypothalamus and project axons throughout the neocortex. Norephinephrine is discharged from A_1 to A_7, dopamine is discharged from A_8 to A_{15}, and epinephrine is discharged from the C system. The B system discharges serotonin. Amines are produced at the cell body and transported through the axon. Functionally, amines affect the mental state of the human, such as emotion[Buck, 1986]. Dopamine, norepinephrine, and epinephrine produces pleasure, anger, and fear, respectively. Several experiments demonstrate effect on memory and learning [Gold, 1984; Hall, 1992].

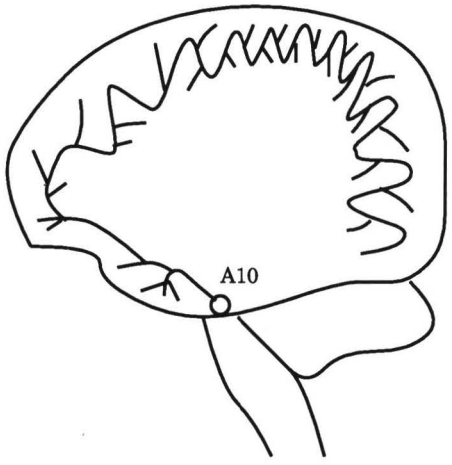

Figure 3: The Projection of A_{10} Neural System

2.3. Neuropeptide

The third class of chemicals are neuropepetides, also a kind of hormone. In the 1960s, de Wied proposed that *adrenocorticotropic hormone* (ACTH) and *vasopressin* (VP) modulate memory and learning by acting on the central nervous system [de Wied, 1965; 1974]. Research indicated that memory is both enhanced and impaired by post-training treatment affecting the monoaminergic[Gold, 1984], cholinergic [Flood and Cherkin, 1986], and inhibitory amino acid systems [Castellano and Pavone, 1988] as well as peptide systems, including ACTH[de Wied, 1974], VP [de Wied, 1984], and opioid peptide [Castellano, 1975]. Recent findings indicate that other peptides, including substance P [Schlesinger et al., 1986], CCK [Flood et al., 1987b], angiotensin II [Yonkov et al., 1986], somatostatin [Vecsei et al., 1986], and neuropeptide Y [Flood et al., 1987a], modulate memory storage.

In general, neuropeptides mediate slow transmissions. Currently, over 40 neuropeptides have been identified in the central nervous system, and several neuropeptides were found to co-exist at one pre-synaptic site.

3. Implications for Learning Theory

Models of neural networks and learning theories must be revised in order to incorporate functions of biogenic amines and neuropeptides. Due to the limitations of paper length, we focus on a model incorporating functions of biogenic amines. Among many possible impacts, this section focuses on three aspects that have immediate impacts on learning theory.

3.1. Elevation of Action Potential

With regard to the forward propagation of a signal, the current model uses equation:

$$a_i = \sum_{j=0}^{N} w_{ij} a_j - h \qquad (1)$$

where a_i is the activation level of the neuron i, w_{ij} is a weight between neuron i and j, and h is the bias. By incorporating hormonal activity, this needs to be revised to:

$$a_i = \sum_{j=0}^{N} w_{ij} a_j + \alpha \sum_{k}^{M} h_{ik} - \beta \quad (2)$$

$$h_{ik} = \gamma R_{ik} C_{ik} \quad (3)$$

$$C_{ik} = \int_0^T d_{ik} a_{kt} e^{-\mu(T-t)} dt \quad (4)$$

where α, β and γ are constants, a_{kt} is the activation level of a neuron discharging the chemical messenger k at time t, and R_{ik} is an amount of receptors for the chemical messenger k on synapse i. R_{ik} may be a constant, or it may change depends on stimuli and genetic information. With higher discharge of amines, the base level to trigger activation of the neuron i decreases. Thus, the circuit consisting of the neuron with the receptor will be more sensitive to the input.

3.2. Modulation of Weight Modification

Next, amines also affect learning in neural networks. Take the example of hebbian learning, the current model uses the equation:

$$\Delta w_{ij} = \varepsilon a_i a_j \quad (5)$$

where ε is a constant, a_i and a_j are activation levels of neurons i and j, respectively. This needs to be changed to:

$$\Delta w_{ij} = \sum_{k}^{M} \zeta_k h_{jk} a_i a_j \quad (6)$$

where ζ_k is a parameter defining positive or negative effects of hormone k for weight up-date strength. This equation means that the degree of weight modification is modulated by hormonal discharge around the synaptic site.

3.3. Mental Rehearsals

Speculating from the elevated action potential of the neurons and known neural circuits in the brain, we may be able to assume the existence of mental rehearsal and its modulation by hormones. Mental rehearsal in this paper means the repeated internal exposure of patterns for memorization. This process does not have to be a conscious process. We speculate that mental rehearsal occurs through the closed loop formed by upward projections from the hippocampus to the neo-cortex and back projections from the neo-cortex to the hippocampus. With a crude abstraction, this can be schmatized as in Fig. 4. Short Term Memory (STM) represents the hippocampus, and three Long Term Memory (LTM) units represent regions of the neo-cortex receiving specific neural projections from the limbic system. There is state-dependency in each region of LTM units. The evaluation unit represents a part of the limbic system which evaluates preference of the given inputs and activates neural systems A, B, and C. Metaphorically, this may correspond to various emotional states.

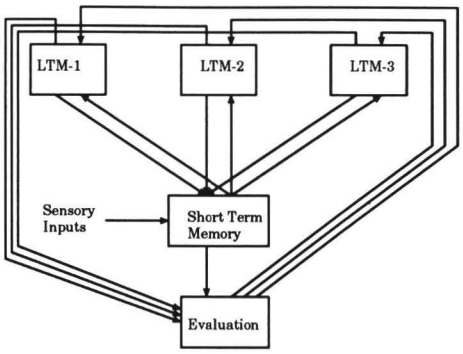

Figure 4: Structure

In this circuit, inputs are quickly learned by STM, and the closed loop between STM and LTM creates a resonance, a repeated propagation of pulses representing certain patterns. Without an active hormonal discharge, this resonance resolves very quickly. However, with an active hormonal discharge, i.e. active A, B, and C neural systems, the baseline action potential for neurons in LTMs is elevated so that sustained (or longer duration) resonance can be created. Some reinforcement learning use mental rehearsal to speed-up learning [Sutton, 1990; Lin, 1990], however, this model adds the modulation of the mental rehearsal frequency.

In summary, hormones modulate learning so that data which are evaluated as important are learned with more frequent mental rehearsal and a higher learning rate than unimportant data. This is critical for survival because life-related situations form only a fraction of the events we are exposed to in daily life. Thus the successful agent must evaluate the importance of the event to enhance the learning effect. This conclusion is consistent with psychological studies on emotion [Ortony, 1987]

4. A Model

Discussions in the previous section lead us to propose a new model of reinforcement learning. The model is a variant of reinforcement learning, and can be applied to existing learning algorithms. A basic architecture is shown in Fig. 5. The model consists of a genetically-determined reaction network (`React`), an evaluation network (also genetically determined) (`Eval`), and a policy network (`Policy`).

When an input is given to an agent implementing the model, it is received by `Policy`, `Eval`, and `React`. Given the input, `React` produces motor control reactions, `Eval` produces reinforcement reward (`reward`) and

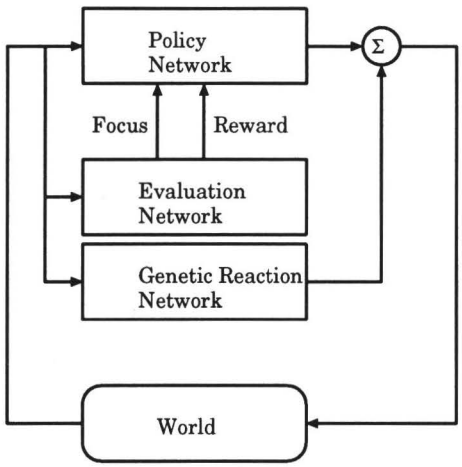

Figure 5: Agent architecture

focus (**focus**) which will be explained later, **Policy** produces a motor control output. The final motor action is determined by adding outputs from **React** and **Policy**.

In the model **React**, **Eval** and **Policy** are implemented using feed-forward neural networks whose initial connection weights are genetically determined by means of the GA process. In this experiment, we used the direct encoding method [Whitley and Hanson, 1989]. For **Policy**, weights are modified through a learning mechanism. **React** and **Policy** use a 16-8-8 network, and **Eval** uses 17-8-2 network. For **Eval**, there are 17 input nodes because 16 for sensory inputs and 1 for internal energy level. Activation levels of two output nodes correspond to **Reward** and **Focus**. In this paper, we use one policy network rather than three networks. As a learning mechanism for the model, we employ a modified version of Evolutionary Reinforcement Learning (ERL: [Ackley and Littman, 1992]). The learning component of the ERL is Complementary Reinforcement Back Propagation (CRBP). Fig. 6 shows the algorithm of CRBP.

We made an extension to the original CRBP to incorporate learning modulation by hormones. CRBP in its original form only takes reward to change its learning behavior, and no modulation is made on the maximum number of mental rehearsals[1] nor the learning rate parameter. In our model, **focus** produced from **Eval** captures the level of hormonal discharge, thus affecting the learning rate and the maximum amount of back propagation allowed for learning. When **focus** is almost zero, almost no mental rehearsal is allowed. A high **focus** level allows a large number of mental rehearsals and a larger learning rate. This means that **focus** level determines which event is important. The existance of **focus** is the major extension from ERL. Fig 7 shows a new algorithm. Revised parts are shown with underlines. **focus** and **reward** correspond to r and f in Fig. 7. With this modification, a revised CRBP should be able to focus learning on significant data.

[1]We use the term *mental rehearsal* referring to the iterative weight update in CRBP. Specifically, from step 2 to step 6.

1. Receive vector i_t. If $t = 0$ go to 7. Otherwise compute reinforcement $r = f(i_t, i_{t-1})$.
2. Generate output errors e_j. If $r > 0$, let $e_j = (o_j - s_j)s_j(1-s_j)$, otherwise let $e_j = (1-o_j-s_j)s_j(1-s_j)$.
3. Back propagate errors.
4. Update weights. $\Delta w_{jk} = \eta e_k s_j$.
5. Forward propagate again to produce new s_j's. Generate temporary output vector o^*.
6. If ($r > 0$ and $o^* \neq o$) or ($r < 0$ and $o^* = o$), go to 2.
7. Set network input to i_t. Forward propagate to produce s_j's.
8. Generate a binary output vector o. Given a uniform random variable ζ and parameter $0 < \nu < 1$,
$$o_j = \begin{cases} 1 & \text{if } \frac{(s_j - \frac{1}{2})}{\nu} + \frac{1}{2} > \xi \\ 0 & \text{otherwise} \end{cases}$$
9. Perform the action associated with o. Let $t = t + 1$. Go to 1.

Figure 6: Complementary Reinforcement Back Propagation

1. Receive vector i_t. If $t = 0$ go to 7. Otherwise compute reinforcement $r = g(i_t, i_{t-1})$, <u>and focus $f = h(i_t)$. Set $\eta = \beta f$, and the maximum mental rehearsal $m_{max} = \gamma f$.</u>
2. Generate output errors e_j. If $r > 0$, let $e_j = (o_j - s_j)s_j(1-s_j)$, otherwise let $e_j = (1-o_j-s_j)s_j(1-s_j)$.
3. Back propagate errors.
4. Update weights. $\Delta w_{jk} = \eta e_k s_j$. <u>$m = m + 1$.</u>
5. Forward propagate again to produce new s_j's. Generate temporary output vector o^*.
6. If (($r > 0$ and $o^* \neq o$) or ($r < 0$ and $o^* = o$)) <u>and ($m < m_{max}$)</u>, go to 2.
7. Set network input to i_t. Forward propagate to produce s_j's.
8. Generate a binary output vector o. Given a uniform random variable ζ and parameter $0 < \nu < 1$,
$$o_j = \begin{cases} 1 & \text{if } \frac{(s_j - \frac{1}{2})}{\nu} + \frac{1}{2} > \xi \\ 0 & \text{otherwise} \end{cases}$$
9. Perform the action associated with o. Let $t = t + 1$. Go to 1.

Figure 7: CRBP with hormonal modulation

5. Experiments

5.1. Artificial world

The experiment assumes an artificial world consisting of a 100 by 100 grid. The emotional agent moves around in this world. There are food sites and enemies. Fig 8

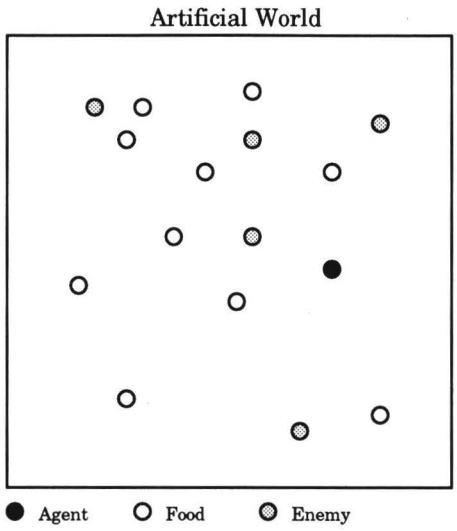

Figure 8: Artificial World

shows a snapshot of the world. Experiments reported in this paper assume 10 food sites and 5 enemies. The agent consumes 1 unit of energy per grid unit moved. The agent will die if its energy reaches zero. Each food site gives the agent 50 units of energy. When the agent is attacked by the enemy, the agent loses 10 points of energy.

Simple sensor systems are available for the agent. It has two types of sensors for 8 directions. Each sensor covers 45 degrees. One type of sensor detects food; the other type detects enemies. Strength of the sensor readings corresponds to the distance to the object.

5.2. Evolution

A genetic algorithm was used to simulate the evolutionary process. The genetic algorithm uses real value direct encoding to represent weights [Whitley and Hanson, 1989], two point crossover, an elitist reproduction strategy combined with proportional reproduction, and an adaptive mutation rate ranging from 1% to 10%. The population size for the following experiments is 20. The initial values for the weights encoded on the chromosome are distributed with a range of ± 2.0. After a predefined time period (100 time units), each individual is evaluated based on its energy level. If an individual has already been killed by an enemy attack, its fitness is equal to the number of time units it has survived.

5.3. Results

A series of experiments was carried out to evaluate the new learning model. The model has been compared with several other models so that four indepdent models were tested, including:

1. Genetic reaction only (`Gen`)
2. Genetic reaction and fixed reinforcement (`Fix`)

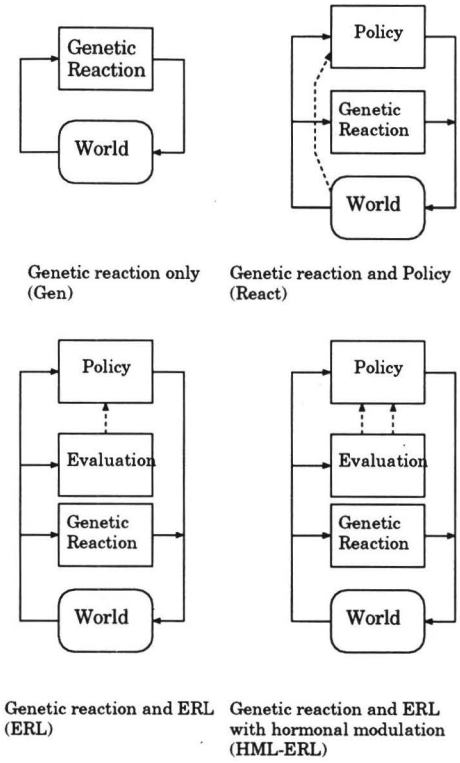

Figure 9: Experimental architectures

3. Genetic reaction and ERL (`ERL`)
4. Genetic reaction and ERL with hormonal modulation (`HML-ERL`)

Fig. 9 shows the architecture for each model. First, an agent with only a genetic reaction circuit `Gen` was evolved and tested in the artificial world. This data is used as a baseline to measure the effects of other circuits on behavior performance. Next, an agent with `Eval` and `Policy` with pre-programmed reinforcement was tested. `Policy` receives positive reinforcement when it eats food, and negative reinforcement when attacked by the enemy. Thus, its learning behavior is based on CRBP with a fixed evaluation circuit. The third experiment was carried out using the model proposed in this paper. The control experiment was carried out using the same architecture, but without modulating the learning rate or numbers of mental rehearsals. The control group is equivalent to evolutionary reinforcement learning (ERL: [Ackley and Littman, 1992]).

The average performance of these variations are shown in Figure 10. This is a typical result from 20 experiments performed on the same task. Several interesting results can be discovered. First, the proposed model attained the best performance. This supports our hypothesis that hormonal modulation improves learning capability. Second, the genetic-reaction-only model (`Gen`) outperformed all other models except the proposed model (`HML-ERL`).

For the model with fixed reinforcement (`Fix`), the performance is significantly worse than the model with ge-

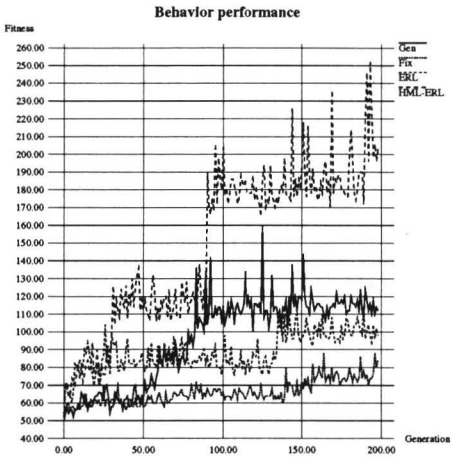

Figure 10: Evolution of behavior performance

netic reaction alone. The problem for this model is that the evolution and behavior of the genetic reaction circuit and the learning and behavior of the policy network are not coordinated, so that two independent behavior modules are evolved and learned without interacting. This experimental result clearly demonstrates that simply adding a learning module on top of the reactive module does not necessarily improve the agent's overall performance.

The surprise comes with the performance of **ERL**. It shows a significantly worse performance than the proposed model **HML-ERL**, and equivalent to or even worse than the agent with only the genetic reaction module (**Gen**). The only difference between the proposed model and this model is in the existence of learning rate and mental rehearsal modulation based on the signals from the evaluation module. In the original ERL, reinforcement learning is triggered for every input ragardless of its importance. Thus, memories of important events are scrambled by weight updates caused by a large number of unimportant and often noisy inputs. The environment in this experiment is sparse than the experiment environment in [Ackley and Littman, 1992] where agents using ERL performed better than genetically-hard-coded agents. Thus, the result is dependent upon the evironment.

6. Discussions

6.1. Evolution of Learning Control

HML is closely related with emotion. Since the discharge of hormone is associated with changes in emotional status, HML can be viewed as emotion-driven learning. Emotion has already been considered as a mechanism of behavior control [Toda, 1985; Pfeifer, 1988]. We consider that emotion also controls learning by regulating the discharge of hormone. However, it is not clear how emotional control emerged through evolution. Experiments indicate that agents with the genetical-determined behavior module, **Gen**, perform almost as well as **HML-ERL**. In some runs, **Gen** outperformed **HML-ERL** at the early stage of evolution. This is because an appropriate evaluation network was not acquired at the early stage of evolution for **HML-ERL**. If several agents co-exist in the same environment, **HML-ERL** and **ERL** could have been extinct due to slow acquisition of an evaluation network. Thus, there is a need to evolve complex agents using an incremental built-up approach. In this case, the first step is to use emotion, more precisely a hormonal discharge, as the sole mechanism of behavior control. The main question is, how was the more complex module can be created and connected to the older module. This is our major future research target.

6.2. Symbol Grounding

HML can be applied as a solution for the symbol grounding problem. The basic approach is to evaluate an input sequence, and decide which input to focus for learning. This means that partitioning of a signal stream may be possible. While various studies demonstrate that formation of attractors from an input stream is possible [Pollack, 1991; Tani, 1995], these input streams need to be partitioned. Each attractor can be viewed as a symbol representing a specific pattern of the sensory input. The input sequences in these studies are either a sequence of symbols [Pollack, 1991], or a sequence of continuous signal with a hand-coded partitioning function [Tani, 1995]. For example, in Tani's robot navigation using dynamical systems approach [Tani, 1995], learning is invoked only when sensory input shows a pattern of a road branching for a maneuver. Without the hand-coded partitioning function, learning is always activated. In this case, attractors are not formed. This suggests that partitioning of the input stream is necessary for the self-organization of the symbols. The question is how the partitioning function emerged. Our hypothesis is that a neural circuit for partitioning, which provides **focus** in HML, was acquired through evolution, and the first level partioning was performed using this circuit. Thus, the acquisition of focusing circuits using hormonal modulation leads to the formation of a primitive set of symbols. If this hypothesis is correct, it means that hormonal modulation plays an important role in symbol grounding.

7. Concludings Remarks

In this paper, we discussed the possible role of hormones in learning, and proposed and examined a new model of reinforcement learning. The proposed model is derived from recent findings in molecular neurobiology. Although the model examined and the task was a simple one, it clearly demonstrates the effectiveness of using hormonal modulation for learning. The result of this experiment has two implications. First, it shows that modeling hormonal activities enables the focus of attention for learning, thereby improving behavioral performance. This would be an important approach for autonomous agents behaving in the real world, because such an agent needs to determine what is important for its survival or task accomplishment. The proposed model enables the

agent to learn vital events even if they are rare events. Seocnd, the result supports biological findings regarding memory modulation by hormones from computational perspectives. While there is much evidence for memory modulation by hormones, these are at either molecular level or behavioral level. No study so far has documented the role of hormonal modulation at the mesoscopic level.

Although we used ERL as a basis of the experiment and the model was proposed as a modification of ERL, the basic idea can be applied to other learning algorithms. Now, it is an open question whether the method proposed has universality so that it improves learning performance when applied to various learning models. If this is the case, the hormonal modulation enables the focusing of learning by evaluating learning data on the basis of the system's goal.

Further research topics will be to incorporate the effects of neuropepetides and multiple amines for focusing.

Acknowledgments

I would like to thank members of Sony Computer Science Laboratory for useful discussions, and Prof. Susumu Tonegawa (MIT) and Dr. Gen Matsumoto (ETL) for insightful comments.

References

[Ackley and Littman, 1992] Ackley, D. and Littman, M., "Interactions Between Learning and Evolution," Langton, C. et al. (Eds.)*Artificial Life II,* Addison-Wesley, 1992.

[Buck, 1986] Buck, M., "The Psychology of emotion," (eds.) LeDoux, J. E. and Hirst, W., *Mind and Brain,* Cambridge University Press, 1986.

[Castellano and Pavone, 1988] Castellano, C. and Pavone, F., "Effects of ethanol on passive avoidance behavior in the mouse: involvement of GABAergic mechanisms," *Pharmacol. Biochem. Behav.* 29, 321-24, 1988.

[Castellano, 1975] Castellano, C., "Effects of morphine and heroin on discrimination learning and consolidation in mice," *Psychopharmacology* 42, 235-42, 1975.

[de Wied, 1965] de Wied, D., "The influence of the posterior and intermediate lobe of the pituitary peptides on the maintenance of a conditioned avoidance response in rats," *Int. J. Neuropharmacol.* 4, 157-167, 1965.

[de Wied, 1974] de Wied, D., "Pituitary-adrenal system hormones and behavior," Schmitt, F. and Worden, F. (eds.) *The Neurosciences: Third Study Program,* The MIT Press, 1974.

[de Wied, 1984] de Wied, D., "Neurohypophyseal hormone influences on learning and memory", Lynch et al. (eds.) *Neurobiology of Learning and Memory,* 1984.

[de Wied and Bohus, 1978] de Wied, D. and Bohus, B., "The modulation of memory processes by vasotocin: The evolutionary oldest neurosecretory principle," *Prog. Brain Res.* 48, 152-156, 1978.

[Flood and Cherkin, 1986] Flood, J. and Cherkin, A., "Scopolamine effects on memory retention in mice: a model of dementia?", *Behav. Neural Biol.* 45, 169-84, 1986.

[Flood et al., 1987a] Flood, J., Hernandez, E. and Morley, J., "Modulation of memory processing by neuropeptide Y," *Brain Res.* 422, 218-34, 1987.

[Flood et al., 1987b] Flood, J., Smith, G. and Morley, J., "Modulation of memory processing by cholecytokinin dependence on the vagus nerve," *Science,* 236, 832-34, 1987.

[Gold, 1984] Gold, P., "Memory modulation: Roles of peripheral catecholamines," Squire, R. and Butter, N. (Eds.)*Neuropsychology of Memory,* Guilford, 1984.

[Hall, 1992] Hall, X., *An Introduction to Molecular Neurobiology,* Sinauer Associates, 1992.

[Lin, 1990] Lin, L., "Self-improving reactive agents: Case studies of reinforcement learning frameworks," *Proc. of Simulation of Adaptive Behavior 1,* 1990.

[McGaugh, 1989] McGaugh, J. L., Involvement of Hormonal and Neuromodulatory Systems in the Regulation of Memory Storage, *Ann. Rev. Neurosci.,* 1989, 12:255-87.

[Ortony, 1987] Ortony, A., "Subjective importance and computational models of emotions," Hamilton, V. el al. (Eds.) *Cognitive perspectives on emotion and motivation,* Kluwer Academic Publisher, 1987.

[Pollack, 1991] Pollack, J., "The Induction of Dynamical Recognizer," *Machine Learning,* Vol. 7, Num. 2/3, 227-252, 1991.

[Pfeifer, 1988] Pfeifer, R., "Artificial intelligence models of emotion," Hamilton, V., Bower, G. H., and Frijda, N. H., (Eds.) *Cognitive perspectives on emotion and motivation,* Kluwer, 1988.

[Schlesinger et al., 1986] Schlesinger, K, et al., "Substance P facilitation of memory: effects in an appetitively motivated learning task," *Behav. Neural Biol.,* 45, 230-39, 1986.

[Sutton, 1990] Sutton, R., "Reinforcement Learning Architecture for Animat," *From Animals to Animats: Proc. of SAB-90,* MIT Press, 1990.

[Tani, 1995] Tani, J., "Self-Organization of Symbolic Processes through Interaction with the Physical World," *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95),* Montreal, 1995.

[Toda, 1985] Toda, M., "Emotion viewed as tightly organized, genetically determined system of behavior-selection programs," Spence, J. & Izard, C., (Eds.) *Motivation, emotion, and personality,* North-Holland, 1985.

[Vecsei et al., 1986] Vecsei, L., et al. "Somatostatin and (d-trp68, d-cys^{14})- somatostatin delay extinction and reverse electroconvulsive shock induced amnesia in rats," *Phychoneuroendocrinology,* 11, 111-15, 1986.

[Whitley and Hanson, 1989] Whitley, D., and Hanson, T., "Optimizing Neural Network Using Faster, More Accurate Genetic Search," *Proc. of ICGA-89,* 1989.

[Yonkov et al., 1986] Yonkov, D. et al., "Participation of Angiotensin II in Learning and Memory," *Methods Find. Exp. Clin. Pharmacol.* 8, 203-6, 1986.