Proceedings of International Conference
on Artificial Intelligence

Adaptive Pattern Recognition with a Self-Organizing Learning Architecture

Jong-Chen Chen
Department of Management Information Systems
National Yunlin Institute of Technology, Touliu, Taiwan
Tel: +5-534-2601 ext. 5332, Fax: +5-531-2077
email: jcchen@mis.yuntech.edu.tw

Abstract

Biological systems have enormous adaptability. We have developed a biologically motivated computer model, called the artificial neuromolecular (ANM) system, that is capable of differentiating patterns and tolerating a certain degree of noise in a self-organizing manner. Two biological features, biological-like structure-function relationship and evolution-friendliness, that facilitate self-organizing learning have been built into the system. With these two important features, the system can be molded to perform coherent functions in a specific task. Three pattern sets were used to test the system, ranging from comparatively dissimilar (randomly generated patterns) to comparatively similar (printed Chinese characters). Each consists of one thousand patterns. Experimental results show that the system is able to achieve a high degree of pattern differentiation and degrade gracefully in the face of increasing noise.

1. Introduction

Adaptability is a common feature in biological systems. The idea of the ANM system is to provide the computer with a representation of the internal world of biological systems. Two biological features, biological-like structure-function relationship and evolution-friendliness, that facilitate self-organizing learning have been built into the system. By structure-function relationship, we mean that the functions and structures of systems are closely related. By evolution-friendliness, we mean the fitness of system structures exhibits an adaptive landscape that allows for evolutionary computation.

The major elements of the system are neurons whose input-output behavior is controlled by significant internal dynamics [13-14, 17-22]. The dynamics are modeled by cellular automata, structured to represent the neuronal cytoskeleton (a subneuronal network found in every neuron). Neurons of this type are referred to as cytoskeletal neurons. They are linked into a multilayer network that abstracts some features of visual circuitry. Multiple copies of these networks are embedded in a memory manipulation system. Neurons with memory manipulation capability are referred to as reference neurons [14-15]. The synaptic connections between a reference neuron and cytoskeletal neurons are facilitated if they fire at the same time. Later firing of the reference neuron will cause all the cytoskeletal neurons controlled by it to fire. The ANM system combines these two types of neurons into a single, closely integrated architecture.

The system is educated to perform desired pattern differentiation tasks by evolutionary algorithms (similar to genetic algorithms). These algorithms act at the intraneuronal level to generate a repertoire of cytoskeletal neurons with different pattern processing capabilities. They also act at the interneuronal level (through the memory manipulation system) to orchestrate different cytoskeletal neurons into groups suitable for performing desired pattern processing tasks.

If a system emphasizes too much on the capability of pattern differentiation, it would be overly sensitive to every bit of patterns, and as a consequence lose the capability of tolerating any noise. However, if overgeneralization occurs, a system will lose its capability of differentiating patterns. The goal of the system is to strike a balance in these two extremes in a self-organizing manner.

2. The architecture

Cytoskeletal neurons and reference neurons comprise the central processing component of the ANM system. The functions of these two types of neurons are complementary and synergistic [1,7,14]. Collections of cytoskeletal neurons transduce signals from receptor neurons into spatiotemporal signals for controlling effector neurons. Reference neurons are used to select appropriate subsets of cytoskeletal neurons, which then control the manner in which input patterns are transduced to output patterns. The overall architecture is illustrated in Fig. 1.

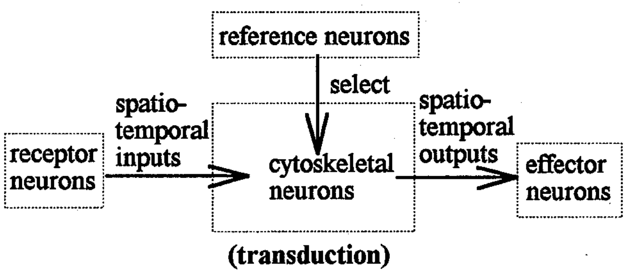

Fig. 1. Overall architecture. Cytoskeletal neurons integrate signals in space and time from receptor neurons into spatiotemporal signals for controlling effector neurons. Only cytoskeletal neurons selected by reference neurons will engage in input/output transduction.

The ANM system as currently implemented comprises eight competing subnets, each consisting of 32 cytoskeletal neurons. Cytoskeletal neurons are manipulated by two levels of reference neurons. Low-level reference neurons select comparable cytoskeletal neurons in each subnet (i.e., neurons that have similar cytoskeletal structures). High-level reference neurons select different combinations of the low-level reference neurons. Fig. 2 provides a simplified picture (only two of the competing subnets were shown, each consisting of only four cytoskeletal neurons). At any instant of time, only cytoskeleton neurons activated by reference neurons will engage in input-output pattern transduction. For example, in Fig. 2, the activation of the high-level reference neuron Ra will activate the cytoskeleton neurons E1 and E4 in each subnet (through the activation of the low-level reference neurons r1 and r4). The activated cytoskeleton neurons will integrate signals from receptor neurons into signals for controlling effector neurons. Evolutionary algorithms act on these corresponding neurons to create the repertoire of neuronal types (i.e., neurons with different pattern processing capabilities) and act at the level of reference neurons to orchestrate these types into coherent groupings (i.e., grouping pattern processing neurons to constitute an effective pattern processing system).

The I/O interface of the system comprises 64 receptor neurons and 32 effector neurons. The connections between cytoskeletal neurons of each competing subnet and its I/O interface are the same. This ensures that corresponding cytoskeletal neurons in each subnet with similar intraneuronal structures will receive the same input from receptor neurons, and that the outputs of the system are the same when the firing patterns of each subnet are the same. Effector neurons are divided into four groups, representing four different behaviors of the system. Each effector neuron is controlled by eight comparable cytoskeletal neurons (i.e., one from each competing subnet) that have similar cytoskeletal structures. An effector neuron fires when one of its controlling cytoskeletal neurons fires.

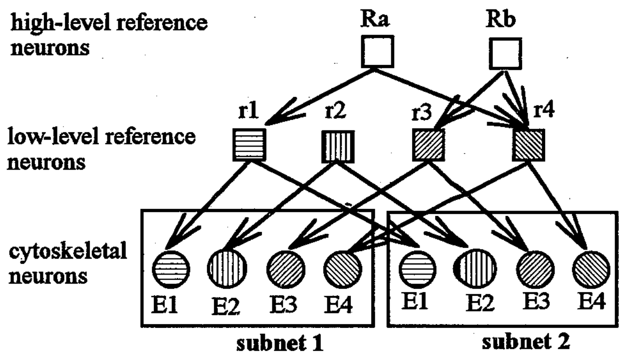

Fig. 2. Connections between reference and cytoskeletal neuron layers. The intraneuronal structures of E1, E2, E3, and E4 in subnet 1 are similar to E1, E2, E3, and E4 in subnet 2, respectively. When Ra fires, it will fire r1 and r4, which in turn causes E1 and E4 in each subnet to fire. Similarly, the firing of Rb will cause r3 and r4 to fire, which in turn fires E3 and E4 in each subnet. (Ei stands for cytoskeletal neuron i.)

3. Pattern processing neurons

Cytoskeletal neurons are the pattern processing neurons in the ANM system. Our implementation of cytoskeletal neurons tries to capture the feature that the cytoskeleton plays the role of signal integration (Fig. 3). That is, they are capable of integrating signals in space and time to yield output signals. The dynamics of cytoskeletal neurons are simulated with 2-D cellular automata [23].

When an external signal impinges on the membrane of a cytoskeletal neuron, it will trigger a unidirectional signal flow along a chain of neighboring components of the same type. For example, in Fig. 3, the activation of the readin enzyme at location (2,2) will trigger a cytoskeletal signal flow along the C2 components of the second column, starting from location (2,2) and running to location (8,2). An activated component will affect the state of its neighboring components of different types when there is a MAP (microtubule associated protein) linking them together. The interactions between two different types of neighboring components are asymmetric. For example, the activation of the C3 component at location (4,8) is not sufficient to activate the C1 component at location (4,7), but stimulates it to a more exciting state. On the contrary, the activation of the C1 component at location (4,7) will activate the C3 component at location (4,8) via the MAP connecting them. The activation of the latter will in turn trigger a signal flow on the eighth column.

Another important feature is that different types of components transmit signals at different speeds and affect each other differently. For example, C1 components transmit signals at the slowest speed, but with the highest activating value. The C3 components transmit signals at the fastest speed, but with the lowest activating value. The activation value of C2 components and their transmitting speed are intermediate between that of C1 and C3 components.

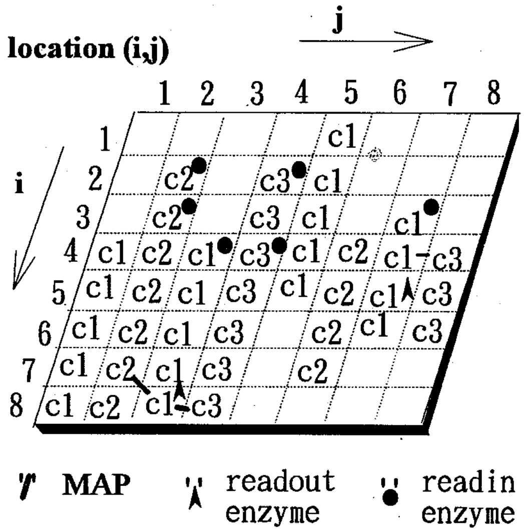

Fig. 3. Cytoskeletal neurons. Each grid location has at most one of three types of components: C1, C2, or C3. Some sites may not have any component at all. Readin enzymes could reside at the same site as any one of the above components. Readout enzymes are only allowed to reside at the site of a C1 component. Each site has eight neighboring sites. The neighbors of an edge site are determined in a wrap-around fashion. Two neighboring components of different types may be linked by a MAP (microtubule associated protein).

When the spatiotemporal combination of cytoskeletal signals arriving at the site of a readout enzyme is suitable, the readout will be activated and then the neuron will fire. For example, there are three possible signal flows that might activate the readout enzyme at location (8,3). The first is a signal flow on the second column, activated either by the readin enzyme at location (2,2) or by the readin enzyme at location (3,2). The second is a signal flow on the third column, activated by the readin enzyme at location (4,3). The third is a signal flow on the fourth column, activated either by the readin enzyme at location (2,4) or by the enzyme at location (4,4). Any two of the above three signal flows might activate the readout enzyme at location (8,3), which in turn will cause the neuron to fire. Nevertheless, the neuron might fire at different times in response to different signal flows along these fibers. One reason is that different types of components transmit signals at different speeds. Another is that signals initiated by different readin enzymes will arrive at a specific readout enzyme at different times.

4. Learning scheme

Five levels of evolutionary variation are possible in the system: at the level of readin enzymes, at the level of readout enzymes, at the level of MAPs, at the level of cytoskeletal components, and at the level of reference neurons. The sequence of evolutionary learning operations of the system is shown in Fig. 4. Evolutionary learning at cytoskeletal neuron and at reference neuron levels is shown in Figs. 5 and 6, respectively. All five levels can evolve simultaneously. But in the present implementation, we allow variation-selection operators to act on only one level at a time. That is, one level (or aspect) is open to evolution for a definite number of generations. During this time all the other levels are held constant. (Only parameters at a level are allowed to change whereas parameters at other levels remain the same.) This multiphasic approach appears to facilitate the evolution process [5-6].

repeat
  evolve **reference neurons** for 16 cycles
  evolve **the pattern of readin enzymes** for 16 cycles
  evolve **reference neurons** for 16 cycles
  evolve **cytoskeletal components** for 16 cycles
  evolve **reference neurons** for 16 cycles
  evolve **the pattern of MAPs** for 16 cycles
  evolve **reference neurons** for 16 cycles
  evolve **the pattern of readout enzymes** for 16 cycles
until (learning objective completed) or
  (maximum learning time reached)

Fig. 4. Pseudocode description of learning scheme.

5. Experimental results

Previous experimental results have shown that the combination of reference neurons and cytoskeletal neurons yields significant computational and learning synergies [2,5-7]. They have also shown that evolution friendliness increases as the number of types of components in the cytoskeletal neurons is increased [8,11]. Learning speed and pattern categorization rate is controllable through the structure of training set [1,3,9]. The input-output behaviors of the system are modified in a gradual changed manner that facilitates evolutionary learning [4,10-11]. Recently, the system has been applied to network packet routing problem [12].

Three types of experiments focusing on the study of adaptability were performed with the system. The first investigated the long-term self-organizing learning capability of the system. The second tested the noise tolerance capability of the system. The third inquired the

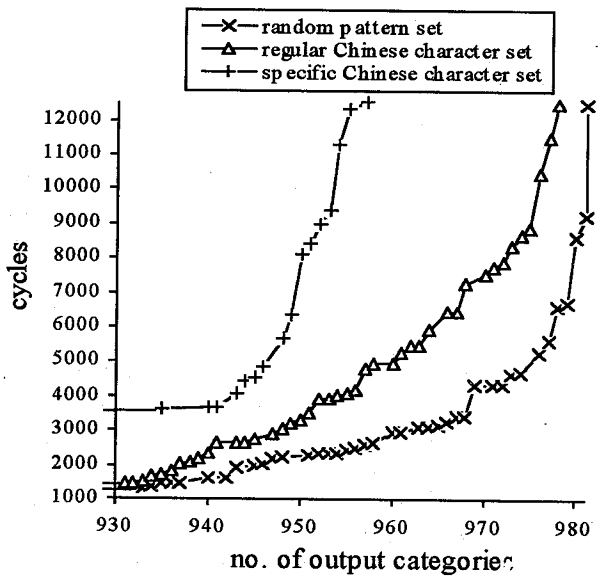

Fig. 8. Effect of similarity of training patterns on learning times. The number of output categories significantly increased in the early stage of learning and slowed down in the later stage. Each mark in a line represents learning proceeds at a specific number of cycle.

Fig. 9 shows that the importance of each bit is different for each training set. For example, in the random pattern set, almost the output of every pattern changes when the second bit in each input pattern is flipped. On the contrary, the outputs of most patterns remain the same when the twelfth bit in each input pattern is flipped. Fig. 9 also shows that different training sets possess different significant bit positions. For example, the importance of location 1 is significantly different in the three training sets. This suggests the system is capable of determining the importance of each bit position, based on the pattern structure of each training set.

5.2 Noise tolerance

Adaptability includes the capability of tolerating noise [16]. The system trained for 8243 learning cycles was tested at 10%, 20%, 30%, 40%, and 50% levels of noise (where these represent the probability that each bit of a training pattern will be changed). At each noise level, ten different test sets were randomly generated from each of the three training sets.

Fig. 10 shows that the ANM system exhibits a high degree of noise tolerance, and that the performance degrades gracefully with increasing noise. The degradation is fastest for the specific Chinese character set and slowest for the random pattern set.

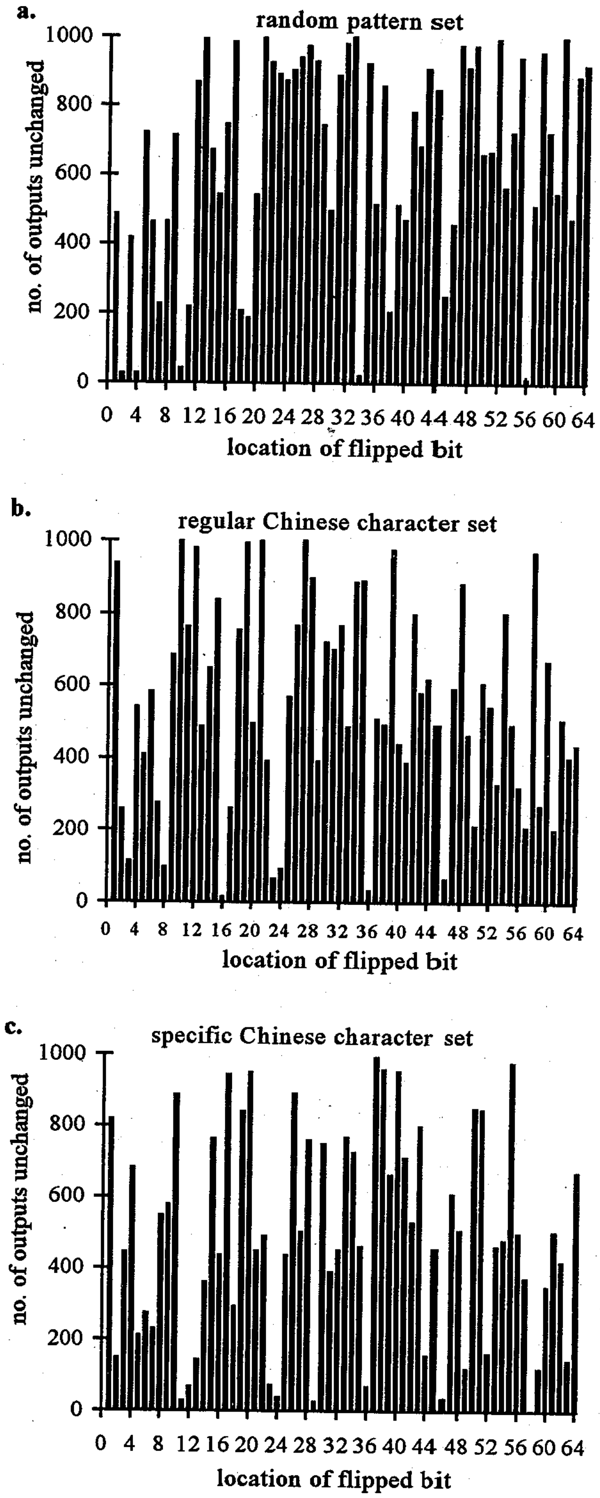

Fig. 9. Effect of all possible single bit changes on similarity of outputs for each input pattern. Each bar represents the number of outputs unchanged when the bit change is introduced. This number changes markedly, indicating that similarity of outputs for each input pattern changes significantly in response to single bit changes.

(a) Regular Chinese character set

(b) Specific Chinese character set

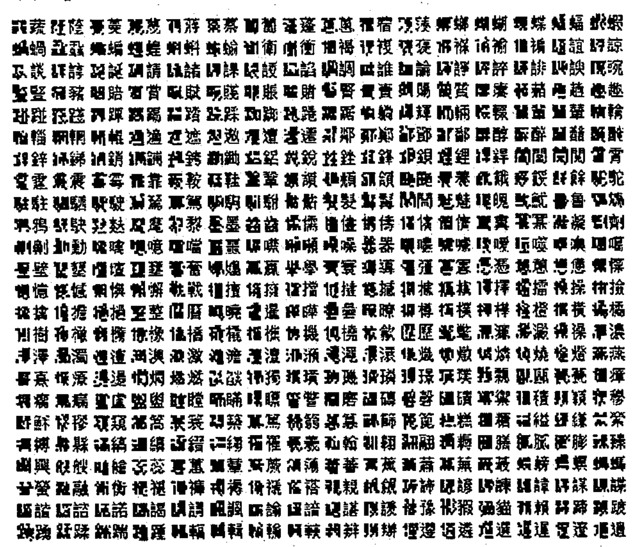

Fig. 7. Chinese character sets. The patterns in the even columns (right) represent 16x15 digitized bit patterns taken from a Chinese character database. These patterns are simplified into 8x8 bit patterns in the odd columns (left), each representing a pattern of receptor neuron activities. Note that the 8x8 bit patterns are adjusted to have the same size as the 16x15 bit pattern for better visualization. The patterns in the specific Chinese character set are much more similar to each other than those of the regular Chinese character set.

5.1 Self-organizing learning

By self-organizing learning, we mean the capability to learn without an external teacher in an indefinite amount of time. Evolutionary learning algorithm (similar to genetic algorithm), a Darwinian variation-selection searching mechanism, can be considered as one of the self-organizing learning algorithms. As mentioned earlier, the system has 32 effector neurons, divided into four groups. For each input pattern, the spatiotemporal sequence of firing effector neurons belonging to different groupings is defined as its output, called an output category. Different such sequences represent different output categories. With four groups of effector neurons in the system, a sequence of seven effector neuron pulses gives 4^7(=16284) possible output categories. The objective is to assign the 1000 patterns in each training set into a maximum number of output categories. The system, in some senses, can thus be considered as a function that maps each input pattern to one of the 16284 output categories. For each training set, the performance improves when the number of output categories mapped by the system increases.

The results have shown that learning is faster when the random pattern set is used than when the regular Chinese character set is used, and that the latter yields yet faster learning than does the specific Chinese set (Fig. 8). The figure also shows the number of output categories significantly increased in the early stage of learning and slowed down in the later stage. Learning is more difficult in the later stage since it has to learn to distinguish those patterns that have not differentiated in the early stage. Our examination of those undifferentiated patterns indicated that they are much more similar to each other.

An important result is that the system shows the continuation of learning during the overall training process (Fig. 8). This is also true for the case of learning in the later stage (note that learning is difficult in the later stage). All of these evidence suggest the system has long-term evolutionary learning capability that allows it to perform pattern differentiation task in a self-organizing manner.

Closer examination of the experimental results indicates that certain bit positions are much more significant than others since they are utilized by the system to differentiate patterns. As a consequence, altering the values of these bit positions will dramatically affect the number of patterns recognized by the system. By contrast, some are less important that the system will maintain a high degree of recognition rate when these bits are altered (from the other view point, they are important since they play the role of tolerating noise).

The following experiment is to investigate which are significant and which are less significant bit positions. It is implemented by systematically making single alternations at each bit position in each pattern of a training set. A first test set was generated by altering the first bit of each training pattern, a second test by altering the second bit, and so forth. In total, this yields 64 test sets. Each comprises 1000 test patterns.

relearning capability after the change of training patterns.

Three training sets were used to test the system, ranging from comparatively similar to comparatively dissimilar. The first set, referred to as the random pattern set, comprised 1000 patterns in which each bit is randomly assigned to be either 0 or 1. Patterns in the second set, referred to as the regular Chinese character set, were randomly taken from a database of frequently used Chinese characters (Fig. 7a). In general, patterns in the regular Chinese character set are much more similar to each other than those in the random pattern set. The patterns in the third set were taken from a database of seldom used Chinese characters, to be called the specific Chinese character set (Fig. 7b). The patterns in the specific Chinese character set are much more similar to each other than those of the regular Chinese character set.

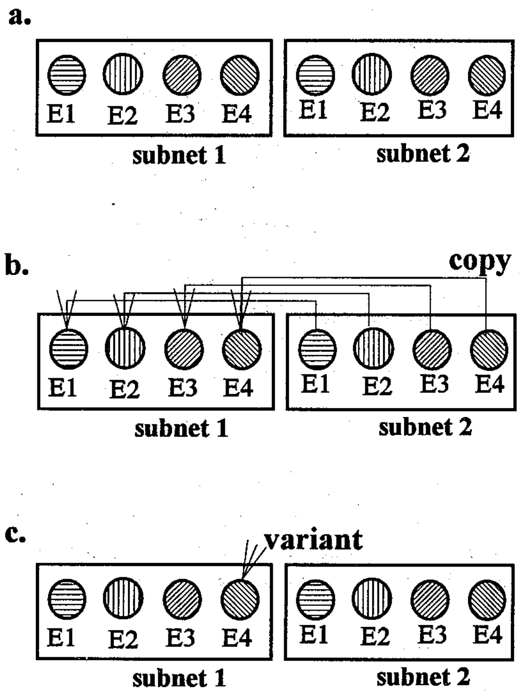

Fig. 5. Evolutionary learning at cytoskeletal neuron level. (a) Each subnet is activated in turn for evaluating its performance. Assume the cytoskeletal neurons in subnet 2 achieve better performance. (b) The pattern of readout enzymes, readin enzymes, MAPs, and other components of subnet 2 is copied to subnet 1, depending on which level of evolution is operative. (c) The pattern of readout enzymes, readin enzymes, MAPs, and other components of subnet 1 is slightly different from that of subnet 2, due to copy error. The figure illustrates a case in which the intraneuronal structures of E4 in subnet 1 undergo variation during the copy process.

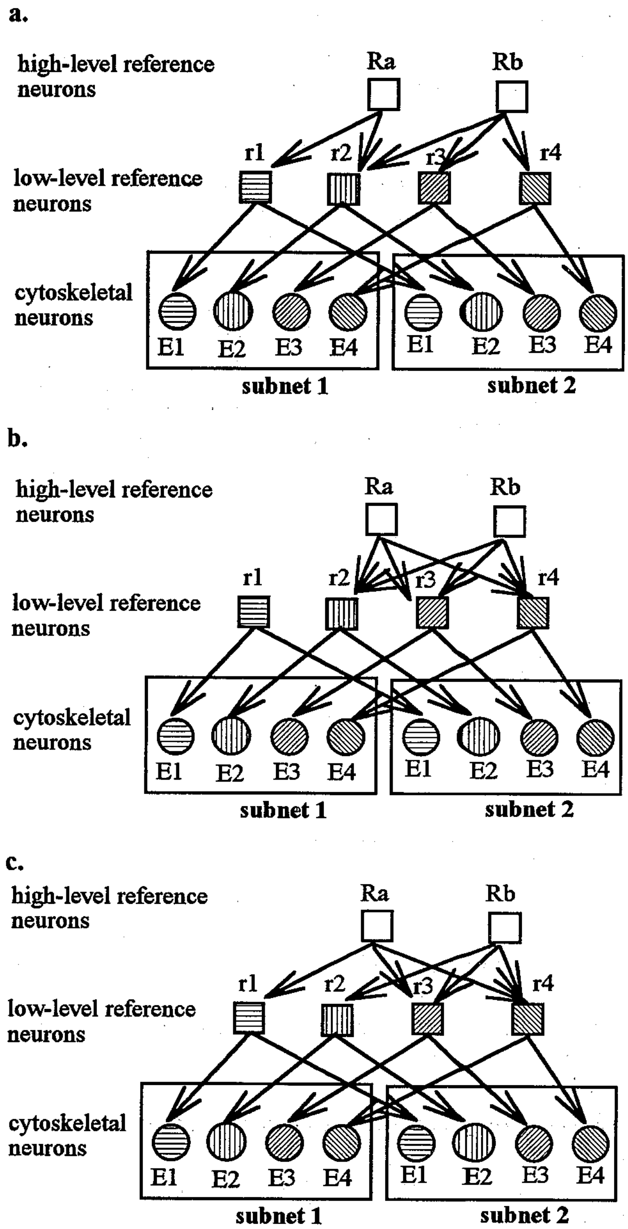

Fig. 6. Evolutionary learning at reference neuron level. (a) Cytoskeletal neurons controlled by each high-level reference neuron are activated (through the low-level reference neurons) in sequence for evaluating their performance. Assume the cytoskeletal neurons controlled by Rb achieve better performance. (b) The pattern of lower-level reference neural activities controlled by Rb is copied to Ra. (c) Ra controls a slight variation of the neural grouping controlled by Rb, assuming some errors occur during the copy process.

When noise is increased to the high noise level of 50%, the number of output categories distinguished by the system decreases from 979 to 912 in the case of the random pattern set (i.e., decreases 6.8%), from 964 to 867 in the regular Chinese character set (10.1%), and from 948 to 827 in the specific Chinese character set (12.8%). The results point to the fact that noise tolerance is less when the training patterns are more similar to each other and greater when they are less similar.

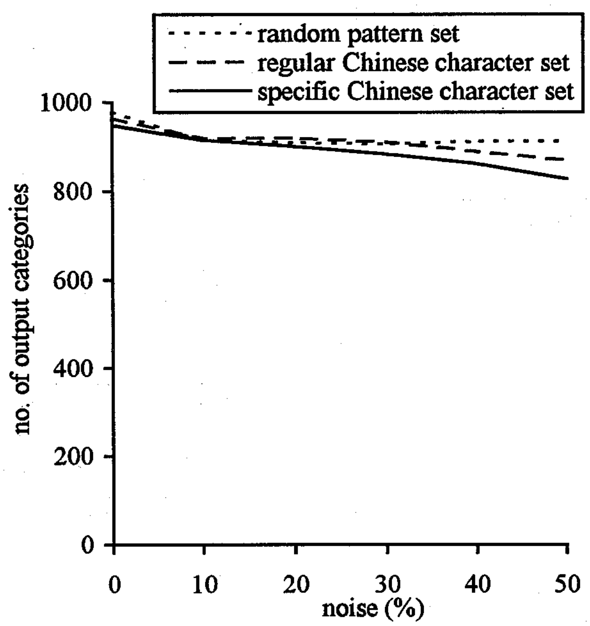

Fig. 10. Dependence of noise tolerance on training set structure. The tolerance decreases most slowly for the random pattern set and most rapidly for the specific Chinese character set.

5.3 Relearning capability (adaptive learning)

The ability to cope with environmental changes is an important feature of biological systems [16]. The experiment to be described below was to investigate relearning capability of a well-trained system (i.e., a system that has been trained with a specific pattern set for a sufficient long time). We first trained the system with patterns in the random pattern set for 8243 cycles (note that the system is able to transduce the 1000 patterns into 979 output categories at this stage). Then, the training patterns were altered, and the system was trained with the modified training set for another 1000 cycles. Note that the modified training sets were generated from the random pattern set by varying a certain number of bits in each pattern, ranging from 1 to 7 bits per pattern. If the system shows continuation of learning after pattern change (i.e., instead of learning from the scratch), this means that it exhibits effective relearning capability subsequent to the modification (i.e., is not trapped by what is previously learned).

As expected, the number of output categories distinguished by the system decreased at the time when the training set was altered. The degradation was much significant when seven bits of each training pattern in average were altered (decreased from 979 to 910 output categories) than when one bit of each training pattern in average was altered (decreased from 979 to 943 output categories). Fig. 11 shows that the pattern recognition rate decreased at the time when the training set was altered, and then the system was able to show continuation of learning after pattern change.

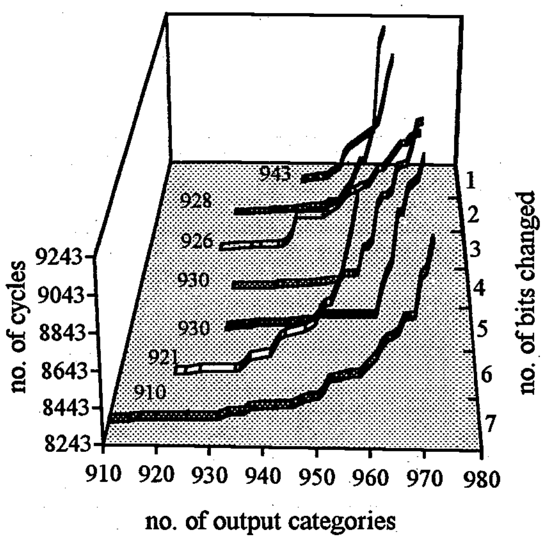

Fig. 11. Effect of varying training patterns on learning times. Each curve represents the learning rate and pattern processing capability when a certain number of bits in each training pattern was altered.

5. Conclusions

We have shown that the system can perform pattern differentiation and tolerate noise in a self-organizing manner, and that learning is more difficult when similar training patterns are used than when dissimilar patterns are used. The noise tolerance experiment demonstrates that the system has a high degree of noise tolerance capability and degrades gracefully in the face of increasing noise. Finally, we demonstrate the system is capable to relearn after the modification of training sets.

Acknowledgment: This research was supported in part by the R.O.C. National Science Council (Grant NSC 86-2314-B-224-001).

References

[1] J.C. Chen, "A Synergetic Multilevel Artificial Brain for Pattern Recognition/Neurocontrol," *Proceedings of the Sixth International Conference on Molecular Electronics and Biocomputing*, Okinawa, Japan, pp. 247-250, 1995a.

[2] J.C. Chen, "Pattern Recognition with Multilevel Learning Algorithms," *Proceedings of the 20th Annual Chinese-American Academic Professional Society Convention*, New York, U.S.A., pp. 88, 1995b.

[3] J.C. Chen, "Simulation of a Discrete-Event Brain-Like Computer Model in Pattern Recognition/Neurocontrol," *Proceedings of the 1996 Summer Computer Simulation Conference*, Portland, OR., U.S.A., pp. 193-198, 1996a.

[4] J.C. Chen, "Pattern Recognition Using an Adaptive Evolutionary Learning Architecture," *Proceedings of the Adaptive, Distributed, and Parallel Computing Symposium*, Dayton, OH., U.S.A., pp. 228-238, 1996b.

[5] J.C. Chen, "A Coevolving Multilevel Learning System for Pattern Categorization and Generalization," To appear in *Proceedings of the First Asia-Pacific Conference on Simulated Evolution and Learning* (in conjunction with Micro-Robot World Cup Soccer Tournament MIROSOT'96), Taejon, South Korea, 1996c.

[6] J.C. Chen and M. Conrad, "Evolutionary Problem Solving with a Multilevel Neuronal Architecture," *Proceedings of the Second Annual Conference on Evolutionary Programming*, San Diego, CA., U.S.A., pp. 127-136, 1993.

[7] J.C. Chen and M. Conrad, "Learning Synergy in a Multilevel Neuronal Architecture," *BioSystems*, 32, pp. 111-142, 1994a.

[8] J.C. Chen and M. Conrad, "A Multilevel Neuromolecular Architecture that Uses the Extradimensional Bypass Principle to Facilitate Evolutionary Learning," *Physica D.*, 75, pp. 417-437, 1994b.

[9] J.C. Chen and M. Conrad, "Pattern Categorization and Generalization with a Virtual Neuromolecular Architecture," To appear in *Neural Network*, 1996a.

[10] J.C. Chen and M. Conrad, "Multilevel Evolutionary Learning in an Artificial Neuromolecular Brain and its Application to the Categorization of Chinese Characters," To appear in *Evolutionary Programming V*, M.I.T. Press, 1996b.

[11] J.C. Chen and M. Conrad, "Evolutionary Learning with a Neuromolecular Architecture: A Biological Motivated Approach to Computational Adaptability," To appear in *Soft Computing Research Journal*, 1997.

[12] J.C. Chen and J.L. Lin, "A Biologically Motivated Multilevel Learning Architecture for Network Packet Routing Problems," To appear in *Proceedings of the Australia-Pacific Forum on Intelligent Processing and Manufacturing of Materials*, Gold Coast, Queensland, Australia, 1997.

[13] M. Conrad, "Evolutionary Learning Circuits," *J. Theoret. Biol.*, 46, pp. 167-188, 1974.

[14] M. Conrad, "Complementary Molecular Models of Learning and Memory," *BioSystems*, 8, pp. 119-138, 1976a.

[15] M. Conrad, "Molecular Information Structures in the Brain," *J. Neurosci. Res.*, 2, pp. 233-254, 1976b.

[16] M. Conrad, *Adaptability*, Plenum Press, New York, 1983.

[17] S.R. Hameroff and R.C. Watt, "Information Processing in Microtubules," *J. Theor. Biol.*, 98, pp. 549-561, 1982.

[18] S.R. Hameroff, *Ultimate Computing*, North-Holland, Amsterdam, 1987.

[19] E.A. Liberman, S.V. Minina, N.E. Shklovsky-Kordy, and M. Conrad, "Microinjection of Cyclic Nucleotides Provides Evidence for a Diffusional Mechanism of Intraneuronal Control," *BioSystems*, 15, pp. 127-132, 1982.

[20] E.A. Liberman, S.V. Minina, O.L. Mjakotina, N.E. Shklovsky-Kordy, and M. Conrad, "Neuron Generator Potentials Evoked by Intracellular Injection of Cyclic Nucleotides and Mechanical Distension," *Brain Res.*, 338, pp. 33-44, 1985.

[21] G. Matsumoto, S. Tsukita and T. Arai, "Organization of the Axonal Cytoskeleton: Differentiation of the Microtubule and Actin Filament Arrays," in *Cell Movement, Vol. 2: Kinesin, Dynein, and Microtubule Dynamics*, F.D. Warner and J.R. McIntosh, eds. Alan R. Liss, New York, pp. 335-356, 1989.

[22] P. Werbos, "The Cytoskeleton: Why it may be Crucial to Human Learning and to Neurocontrol," *Nanobiology*, 1, pp. 75-95, 1992.

[23] S. Wolfram, "Cellular Automata as Models of Complexity," *Nature*, 311, pp. 419-424, 1984.