'nternational Workshop on Biologically Inspired Evolutionary Systems
Tokyo, Japan--May 30-31, 1995

A Primitive Model of Language Generation by Evolution and Learning

Takaya Arita [1], **Keiichi Unno** [2], and **Kimio Kawaguchi** [3]

[1] School of Informatics and Sciences, Nagoya University
Huro-cho, Chikusa-ku, Nagoya, 464-01 JAPAN
ari@info.human.nagoya-u.ac.jp
[2] Yokogawa Digital Computer Corporation
[3] Department of Electrical and Computer Engineering, Nagoya Institute of Technology

Abstract

Natural language, communication or related mental phenomena must surely be a prominent candidate for an evolutionary explanation. This paper discusses a primitive model of language generation by evolution and learning among a population of artificial organisms whose brains are realized by a model of associative memory with a neural network structure. The goal of our study is to acquire general knowledge of the theory that relates the mechanisms to the evolutionary process such as language generation, and to develop the evolutionary systems which have facilities for still more intelligent information processing.

1 Introduction

The so called *synthetic approaches*, which are the engineering approaches to the mechanism concerning information processing in the human brains, have been employed to construct models grounded on physiological evidence and some hypotheses, and to analyze the characteristics of the models. These models not only give informative suggestions to study on the brains, but also are applicable to development of new systems which have facilities for more intelligent and flexible information processing. There are various kinds of these approaches: Some researchers make microscopic models for neural cells and axons. Some give attention to a specific function or specific system of the brain, such as is employed in the visual system. Others seek to make models for the total function based on a complex system. The model described in this paper belongs to the third category.

This paper presents a model which concerns with language generation in a population of organisms, based on the synthetic approaches. There has already been a language generation model to indicate what kind of processing can be accomplished by using a neural network [1][2]. A model of associative memory with a neural network structure is used as the memory in the brain of the artificial organism named *Langy*. Two organisms give specific names to the concepts which are extracted from the stored information, and interchange them with each other. As they repeat common experiences and modify their own words by learning according to what the other Langy says, the differences between the words become smaller and smaller, and finally the two organisms agree on a word for each object. In other words, they learn to exchange information, which is expressed by stimulus patterns, through the medium of uttered words, and to act accordingly on the words uttered by the other organisms [2].

Recently, a new science called *artificial life* has been established, which is considered to be an extension of the synthetic approaches. Its field as a whole represents an attempt to vastly increase the role of synthesis in the study of biological phenomena. Furthermore, in the artificial life approach, we need not restrict ourselves merely to attempting to recreate biological phenomena that originally occured naturally, and we have the entire space of possible biological structures and processes to explore, including those that never did evolve on earth. It is the role of synthesis in artificial life study to give us a glimpse of that wider space of possible biologies [3].

Natural language, communication or related mental phenomena must surely be a prominent candidate for an evolutionary explanation, since we cannot think of it as the results of explicit agreement among humans. From this point of view, we expand the Langy model by introducing the artificial life methodology into it. We present a new model, named *LangE*, and investigates not only generation of language by using a neural network, but also evolution of language by using a framework of the genetic algorithms [4]. In this LangE model, a population of the artificial organisms inhabits a lattice plane, and each LangE repeats "conversations" with neighbors. LangEs can reproduce with the child's genome derived by combining information from two parent genomes, based on a distributed and asynchronous genetic algorithm. This algorithm is essential when attempting to observe the phenomena of language propagation or "dialect" within one environment. This paper discusses generation and growth of language, and seeks for general knowledge of how to construct the evolutionary systems, based on the results of the experiment upon the LangE model.

2 Language Learning

2.1 World image construction

When a human looks and eats an apple, information sensed at receptors, e. g. eyes, ears, nose and mouth, is transformed to corresponding stimulus patterns in the brain (Fig. 1). Furthermore, in hearing the word "apple", this sound is also transformed to a stimulus pattern in the brain. One of the significant functions of the brain is to synthesize and memorize pertinent stimulus patterns while maintaining the appropriate relations among them. Namely, the attributes of apples and the word "apple" are associated and mutually recalled. Therefore, all the attributes of apples can be recalled from a part of them including the word "apple".

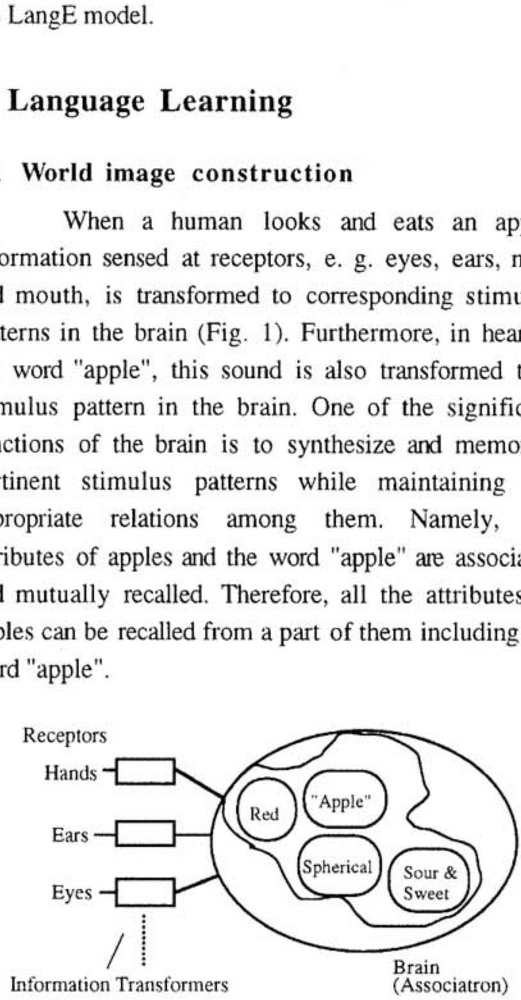

Fig. 1 Transformation of information.

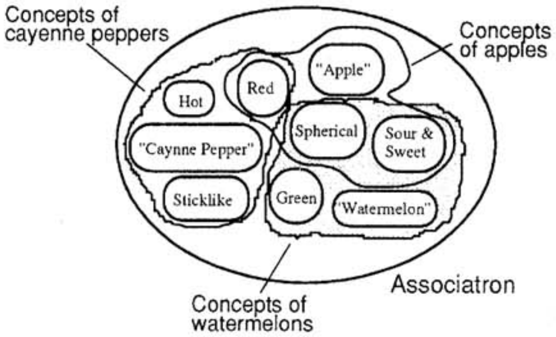

Fig. 2 An aspect of memory in the brain.

When a human observes objects, an exciting pattern, corresponding to attributes of them, arises in the brain, because the brain is self-organized to hold the relations among phenomena in the outer world. The relations stored in the brain memory construct a kind of inner model of the outer world. The authors call this inner model a *world image*. Fig. 2 shows that three objects are stored in the memory.

In the same way, the artificial organisms (Langys) are able to construct such a kind of inner model of the outer world. The brains of Langys are realized by *Associatron* [5], which is a model of associative memory with a simple neural network structure. The principle of Associatron is as follows. Items to be memorized are represented as n-dimensional vectors, whose elements take -1, 0, or 1.

$$\mathbf{x}^{(p)} = (x_1^{(p)}, x_2^{(p)}, \ldots, x_i^{(p)}, \ldots, x_n^{(p)})^t,$$

where p is the index of the items, and t denotes transposition. Items are memorized as the sum of the auto-correlation matrices of the vectors, that is,

$$M = \sum_{p=1}^{k} \mathbf{x}^{(p)} \cdot \mathbf{x}^{(p)\,t},$$

which corresponds to the Hebbian rule [6].
Memorized vectors are recalled by

$$\mathbf{z} = \mathrm{sgn}(\mathrm{sgn}(M) \cdot \mathbf{x})$$

where sgn(u) is a threshold function, defined as

$$\mathrm{sgn}(u) = \begin{cases} -1 & \text{if } u < 0, \\ 0 & \text{if } u = 0, \\ 1 & \text{if } u > 0. \end{cases}$$

In the case that this function is applied for matrices or vectors, the above operation is carried out for each element of matrices or vectors. Recalling process defined by

$$\mathbf{z} = \mathrm{sgn}(M \cdot \mathbf{x})$$

makes no difference essentially.

If most elements of input vector $\mathbf{x}$ are equal to the corresponding elements of $\mathbf{x}^{(r)}$, and the rests are 0s, then it is expected that the recalled vector $\mathbf{z}$ is equal or similar to $\mathbf{x}^{(r)}$. This means that Associatron can recall the entire memorized pattern from only a part of it.

This function can be implemented by using the mutually-connected neural network structure (Fig. 3). Associatron is composed of neurons which correspond to the elements of item vector. Individual pairs of neurons are connected each other. When an input pattern is fed in and the excitation pattern arises in this neural network, the synaptic weights are increased by the products of input values of neurons on both sides of the synapse. At the same time, individual neurons stimulate other neurons through synaptic connection. The stimulation strength is the product of the output from the neuron and the quantized value of synaptic weight (1, 0 or -1, according to the value, plus, zero or minus, respectively). Each neuron takes a 1, 0 or -1 value by majority decision, that is, it takes 1, if there are more stimuli 1s than -1s, and it takes -1, if there are more -1s than 1s.

In the case that input part and output part of the system are separated from each other, it is more effective to use cross-correlation instead of using auto-correlation. Operations of this kind of Associatron are as follow: The memorizing process is

$$M = \sum_{p=1}^{k} \mathbf{y}^{(p)} \cdot \mathbf{x}^{(p)t},$$

and the recalling process is

$$\mathbf{z} = \mathrm{sgn}(\mathrm{sgn}(M) \cdot \mathbf{x})$$

or

$$\mathbf{z} = \mathrm{sgn}(M \cdot \mathbf{x}),$$

where $\mathbf{x}$ and $\mathbf{y}$ express input vector and output vector, respectively. Inversely, it is also possible to recall $\mathbf{x}$ from $\mathbf{y}$ by

$$\mathbf{z} = \mathrm{sgn}(\mathbf{y}^{t} \cdot \mathrm{sgn}(M))$$

or

$$\mathbf{z} = \mathrm{sgn}(\mathbf{y}^{t} \cdot M).$$

This kind of Associatron is used for memory function in the models described in Section 2 and Section 3.

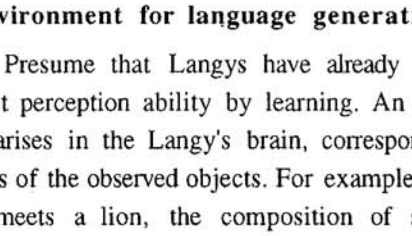

Fig. 3 A neural network for associative memory

2.2 Environment for language generation

Presume that Langys have already obtained sufficient perception ability by learning. An exciting pattern arises in the Langy's brain, corresponding to attributes of the observed objects. For example, when a Langy meets a lion, the composition of stimulus patterns which correspond to the attributes of the lion, e.g. "brown", "with hair", and "big", arises in its brain, based on the idea described in 2.1. When the Langy encounters a rabbit, the composition of stimulus patterns of attributes, such as "white", "with hair" and "with long ears" arises in the same way. Now, presume that, for the same attribute, the same part of the stimulus pattern is active. For example, the attribute "with hair" raises the same pattern in the same position of the stimulus pattern, upon seeing a lion and a rabbit.

Under these assumptions, by taking the intersections of the stimulus patterns, it is possible to extract the attribute "with hair". Since the memory in Langy's brain is implemented by Associatron, this operation is available through the *random stimulation* to Associatron, which is remarkably utilized in the formation of the world image in the Langy model [1][2]. Objects and attributes extracted by random stimulation are treated in the same manner and memorized as *concepts*. At the same time, *names* are given to them. Each Langy carries out these operations independently, and forms a world image of its own. However, it is expected that the world images of the two Langys are alike, because they have similar sensors and receive the same stimuli from the same world.

It seems, from above argument, that attributes have been defined a priori, and then Langys only give one-to-one correspondences between the exciting patterns and names. Considering the following example, however, it is obvious that it is not the case. Assume that a lion has attributes "brown", "with hair" and "dangerous", and that a bear has attributes "black", "with hair" and "dangerous". Then, the common attributes for the two objects are "dangerous and with hair". This is regarded as one concept (a *composite attribute*), because if no other objects were to exist in the world, the attribute "dangerous" and the attribute "with hair" would never be discriminated between. In this sense, so far called "attributes" may as well be called "the minimum units of attributes".

When two Langys recognize an object, they recall the concepts in relation to the object expressed by the stimulus patterns in their world images. Then, they recall words from the concepts and speak them to each other. Only words corresponding to concepts are treated with, since it is too difficult to deal with language including a complicated syntax or context at one bound. Initially, as the words are determined randomly and independently, words for the same concepts are different between two Langys. In having *common experience*, each Langy feeds the association of the words which the other Langy says and the concepts which it recalls for itself, into the Associatron memory. By repeating this operation, the words of two Langys corresponding to the same objects or the same attributes become gradually similar, until finally become identical. Then when one Langy hears the words the other Langy utters, it can recall the same concepts in its brain that the other recalls.

At this stage, so to speak, they share their world images through the medium of words. In this situation, it can safely said that a primitive language has been generated between two Langys. These scenarios are shown and verified by the computer simulations [2].

3 The LangE model

3.1 Environment for language generation

We enlarge the scope of the model described in Section 2, and construct a new model (*LangE*) of language evolution with a virtual world, in which many LangEs reside and repeatedly produce offspring. LangEs, which have the same abilities that the Langys have, are placed in a two-dimensional toroidal grid (Fig. 4), and cannot move out of their location. Every location contains one LangE.

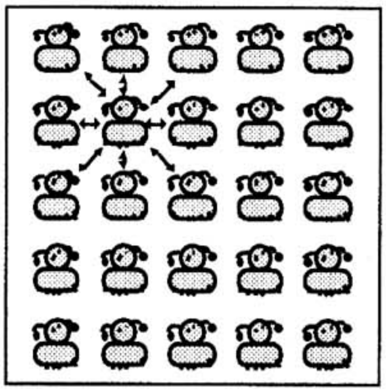

Fig. 4 A population of LangEs

Each LangE is able to construct its own world image by Associatron with the mutually connected neural network structure also in this model. Stimulus patterns which is fed into Associatron, consists of both the patterns for a object and the names given to the objects or concepts.

Each LangE Lij (i and j denote the location in the grid plane) has, as its internal values indicating its state, the scores Pij, the age Aij, the inborn value GAij related to the selection pressure, and the cross-correlation matrix Mij which is the phenotype of and initially equals to the inborn matrix GMij.

3.2 Conversation

A LangE in some location and the surrounding eight neighbors have a common experience by observing an object, recalling the words for it, speaking them to each other, and feeding into the Associatron memory the association of the words spoken by other LangEs and the patterns for the object, as is described in 2.2. Only one-to-one conversations are held between the LangE in the location of the object and the eight neighbors. In the case that the words spoken by the other LangE matches the words the LangE speaks, both are rewarded with some value g, which are added to the scores of them. In the case on the contrary, g is subtracted from the scores of both.

Rewarding the LangEs according to whether they recall the same words which correspond to the patterns representing the object in the world images, has following meanings. Under situation of common experiences, the conversations with such rewarding produce the selection pressure toward the unity of the names, that is to say, generation of a primitive language. Meanwhile, the established conversation under the situation of *sole experiences,* which denotes that an object is observed only one LangE, means information transmission through the medium of uttered words, and makes the other LangE recall the stimulus pattern corresponding to the same object which it doesn't observe.

For example, if the words expressing lions are coincident, when a LangE hears the word "Lion!", it can recall an exciting pattern containing the pattern of the attribute "dangerous", which may make the LangE run away from and survive the lion. Or, if a LangE hears the word "Apple!", it may recall an exciting pattern containing the pattern of the attribute "delicious", which may make the LangE go to take it and escape from starvation. Therefore, the more score Pij a LangE has, the more chances to produce offspring and the less chances to die.

For the sake of simplicity, rewarding depends on whether the words uttered by two LangEs are coincident or not, instead of judging accuracy of the pattern which is recalled when hearing the word of the other LangE. The process of extracting attributes are not analyzed in this paper, since we concentrate ourselves on the analysis of evolutionary process related to language. Therefore, names are given only to the objects that are observed, and only the name given to the observed object is spoken to the other.

3.3 Alternation of generations

Alternation of generations in the LangE world is based on the genetic algorithms [4]. The simple genetic algorithm is composed of initial population generation and repeatedly-executed operators: reproduction, crossover and mutation. Our model adopts a distributed and asynchronous extension of the typical genetic algorithms, because it is unnatural that all the LangEs are compared and some of them are reproduced at fixed time intervals. Each LangE in the population has a distinct genome, of which each gene has an integer value corresponds to GMij(x, y) (an x, y element of the cross-correlation matrix GMij), or GAij which is a positive value related to the selection pressure.

The probability of a LangE to die is,

$$\text{Select}(Aij, Pij, GAij) = 1 / \{ 1 + \exp (-s*((Aij - c*Pij) / GAij - 1)) \},$$

which is calculated on each LangE. independently at every turn-around (c and s are positive constants). If a LangE dies according to the above function and removed from the grid, then offspring is produced at the location, thus keeping the overall population size stable. First, a pair of LangEs are selected from the eight neighbors with the probability in proportion to following weights:

$$\text{Parent}(Pij) = \Theta(Pij)^2 \;/\; \sum_{|i-k| <= 1,\ |j-l| <= 1} \Theta(Pkl)^2,$$

where $\Theta(x)$ equals to p (if $p>0$) or 0 (if $p \leqq 0$).

Mated genome strings cross over at randomly-selected crossing sites. Information not on the learned matrices Mij, but on the inborn matrices GMij is, thus, inherited to offspring. The last operator, mutation, is performed on a bit-by-bit basis, with some probability. The age Aij and the score Pij of the offspring are set to be zero when they are born.

The outlines of evolution processes in the model are shown in Fig. 5. First, initial random population is created in the environment. A location which is a center of conversations and an object which is observed by nine LangEs are randomly selected. Then, they hold eight one-to-one conversations (between the center LangE and surrounding LangEs), which cause learning by the neural networks. Again, another location and a object are selected, and the conversations are held. In this way, every lattice point is selected one time.

Next, if there is a LangE which is to die and removed according to the probability Select(), an offspring is to be created at the location of the removed LangE. A pair of LangEs are selected from the eight neighbors using the function Parent(). Mated genome strings cross over, and mutation is performed with some probability. The processes composed of such conversations and generation alternation are repeated. The passage of time for one repetition considered to be one year.

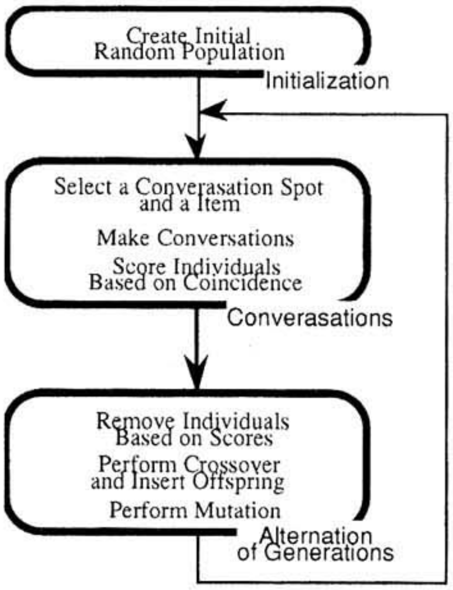

Fig. 5 Main steps of the LangE algorithm.

4 Preliminary Experiments

4.1 Conditions and measurements

Preliminary experiments on the model presented in the previous section are shown in this section under the following conditions. The size of the environment is 5 x 5, in which 25 LangEs live. The length of the input/output field of each Associatron is 55 bits, which consists of the field with the length of 50 bits for the object patterns, and the field with the length of 5 bits for the words corresponding to the objects, as shown in Fig. 6. The number of the objects is six (Fig. 7), the reward or penalty g is 1, and mutation probability e is 1/10. In Select(), c = 1 and s =5. GAij and the elements of GMij for the initial population are given randomly within $100 \leqq GAij \leqq 300$, $-5 \leqq GMij(x,y) \leqq 5$, respectively.

Following six measurements are defined for analyzing the results of experiments. Each measurement is calculated as an average value of all LangEs for every 100 years.

(1) Life-Expectancy:

The average age of LangE when it dies.

(2) Selection Pressure (=GAij):

Inherited parameter used in Selection(). The more, the less likely to die.

(3) Learn Length:

The average age of LangE, when each name given by the LangE and the names given by more than three neighbors become coincident, that is, the average age when LangE shared a *vocabulary* with more than three neighbors.

(4) Genome Level:

The average number of surrounding LangEs which have the same vocabulary as that of the newborn LangE.

(5) Language Coincidence:

Average rewarded score per year. Maximum is 16, when 8 as a center and 8 times 1 as a neighbor, while minimum is -16.

(6) Language Transition:

Number of times of reconstruction of another language unity which is preceded by destruction of language unity. Language unity means that all of the LangEs share identical vocabulary, that is, each object is given its own unique name by all LangEs. The situation that the Language Transition doesn't converge to zero means that the language changes continuously.

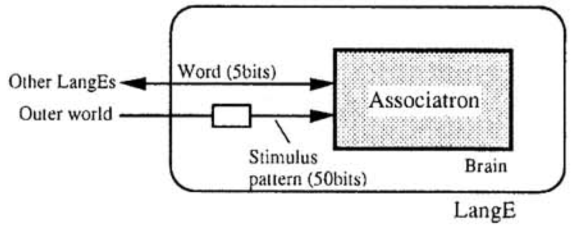

Fig. 6 The brain of LangE.

```
I1:   +-+++----+.....+---+-+++++++++---+-.....-++--++-++
I2:   ++-++..........-+-++...............+-+------+.....
I3:   ..........--++-++----++--.....+++--.....-+++--+++-
I4:   .....+++++.........................+++++.....-+-+-
I5:   +-+-+.....-++-+..........--+--+-+-+---+++--++.....
I6:   +++---+-+++-+--.....---+-+++++++--+.....++++-+--++
```

Fig. 7 Stimulus patterns of objects,

('+' =1, '-' = -1, '.' = 0).

4.2 Evolutionary dynamics

The results are shown in Fig. 8 and Fig. 9. Language Coincidence shows a tendency to increase, in other words, the scores which LangEs acquire are increased by evolution, as might have been expected. Life Expectancy and Selection Pressure also increase at initial stage, and then show a tendency of converging. Genome Level has anti-correlation with Learn Length, because it can be considered, in general, that the more the Genome Level is, the less necessity to learn is. This tendency is shown in these figures. It is also shown that Genome Level has correlation with Language Coincidence. Learn Length shows a peak once, and then decreases. The cause of the decrease seems to be that it becomes easier to learn, according to unity of language. It is a noticeable fact that Language Transition doesn't converge to zero, owing to rapid alternation of generations.

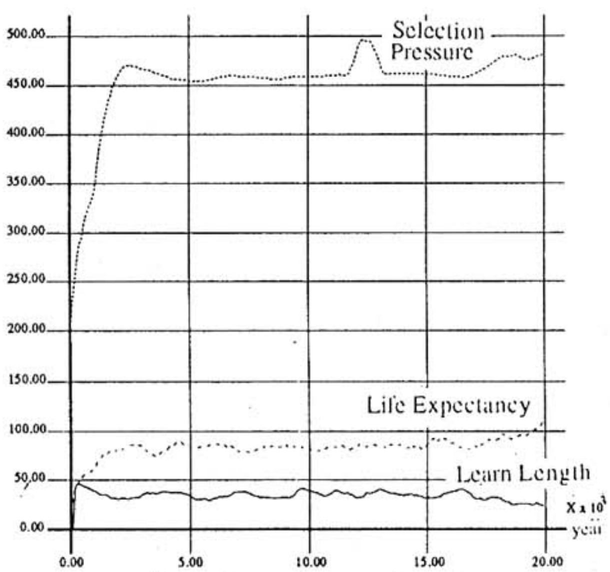

Fig. 8 Simulation results (1),
(Learn Length, Life Expectancy, Selection Pressure).

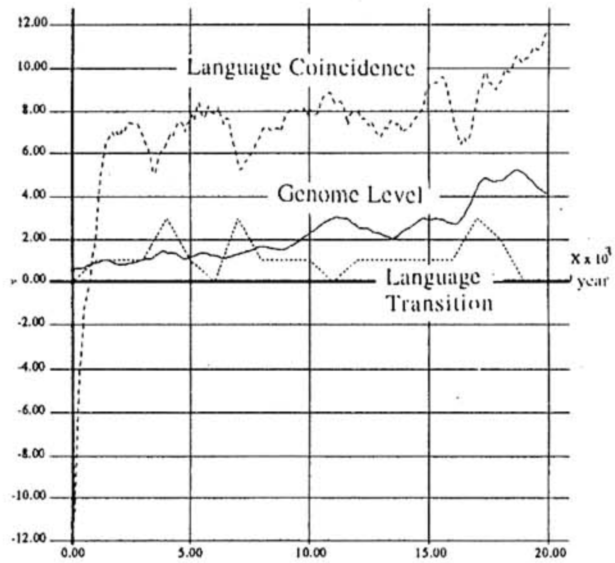

Fig. 9 Simulation results (2),
(Learn Length, Life Expectancy, Selection Pressure).

Fig. 10 shows three typical propagation patterns, each of which shows the transition of the names for some object given by 25 LangEs. Each name is represented by an alphabet for simplicity, which is actually an 5-bit data in the experiments.

Fig. 10a) shows a gradual convergence to "L", which is a typical process from initial state to equilibrium by learning and evolution. Fig. 10c) shows the change from "e" to "E". Such a process as shown in this figure is seen, if there is large probability of dying. Under such situations, frequent occurrence of birth at concentrated location makes the spread of new words, which keeps Language Transition large .

Fig. 10b) shows a unique case of propagation patterns of words. First, the word "J" is diffused among all LangEs. Then, another word "K" rises and spreads, finally generate a remarkable state that "j", which can be called *dialect*, is split by "K". Accumulation of mutation, propagation delay and the effects of matrices given to offspring produce such complex dynamics, while learning by Associatron and selection of parents have large effects on language unify.

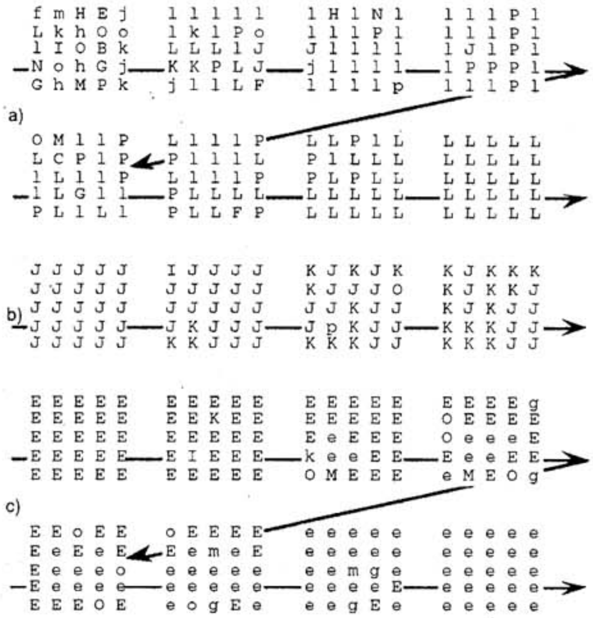

Fig. 10 Examples of name transition,
a) unification of names,
b) split "dialects",
c) transition to another name.

5 Comparison

The LangE model is an enhancement of the model which was constructed for the purpose of indicating the possibility of communication to arise by self-organization using a neural network. We have introduced the artificial life methodology into the model to investigate the evolution of communication among a population of artificial organisms.

A few models concerning evolution of communication in artificial organisms have been proposed in the context of artificial life. It is significant to compare our model with them, since the goal of seeking to evolution of communication is shared.

MacLennan [7] proposes a model, in which the genotype of each simulated organism is represented by a transition table and likewise its phenotype. There is shared and global environment of symbols, and each organism can match and post a symbol to the global environment, based on its transition table. Whenever an organism's action matches that of the most recent symbol posted, both matching organisms receive a credit.

Werner and Dyer [8] explore the evolution of simple communication protocols for mate finding. Female animals have the ability to see males and to emit sounds, and male animals are blind, but can hear signals from females in their model. Simulation resulted in a progression of generations that exhibit effective mate-finding strategies.

In MacLennan's model, inputs from outer world are assumed to be symbolized already as state numbers, and the learning ability of his organisms is attained only by rewriting one line in the transition table, not by self-organizing neural networks as in our model. In Werner and Dyer's model, the brains of the organisms are composed by the neural networks, but have no ability of learning at all.

It is easier and simpler to analyze the general action patterns of the organisms, and to give various meanings to the results of the experiments in our model and MacLennan's model, because abstract actions or environments are processed in general forms. On the other hand, Werner and Dyer's simulation incorporates actual tasks such as moving or mating that is solved using signaling from females to males, and resulted in a progression of generations with increasingly effective strategies.

There is no possibility in MacLennan's model to analyze such a kind of phenomena of language propagation as shown in Fig. 10, because all the organisms share a global environment of symbols, and produce signals in turn on the same conditions. In Werner and Dyer's model, offspring is produced at random location in the environment, though local operators for selection and mutation are adopted in the framework of the distributed and asynchronous genetic algorithm like our model. From this point, phenomena of language propagation or locality of communication are considered not to belong to the main themes to analyze in their model.

Recently, an interesting model of evolution of communication has been proposed by Ackley and Littman [9]. The environment contains three organization levels (individual, local and global), which are interdependent in various ways. An individual's brain is a neural network containing a total of 32 linear threshold units. Various concepts concerning the environment and the behaviors of the organisms are built into the model, and it becomes considerably complicated like many models which investigate ecologies in virtual worlds. It is reported that effective communication based on the exchange of initially arbitrary signals can evolve and stabilize even when it provides no benefit to the individual speaker.

6 Conclusion and Future Directions

Generation and growth of a primitive language by learning and evolution has been discussed in this paper. A new model has been constructed, in which a population of artificial organisms inhabits a lattice plane and each repeats communicating information with neighbors by uttering words. The model has been implemented and we have analyzed its evolutionary dynamics based on the results. The remarkable phenomena of language propagation has been also observed in the simulation.

We are now analyzing the results of experiments furthermore, especially the relation between the parameters and the dynamics of the model. A part of the results of this analysis is described in another paper [10] to throw a side light on our model.

Our model could be extended in several directions. One obvious direction would be to consider more complex world in which more interesting behaviors are allowed. For example, if we incorporate a natural task that can be solved using communication, as opposed to the abstract environments and actions of our model, we would expect emergence of some concrete behaviors. It would be then important to avoid introducing a high degree of arbitrariness in the model.

Another direction would be to evolve complex protocols such as those requiring syntax. One of the ways to get a clue to this would be to utilize the potential of the Associatron model with *total-activity control* [11]. This enhanced Associatron can memorize and recall sequential patterns, that is similar to the so called *reverberation* in the brains of living organisms, by total-activity control.

It is considered to be a promising line of research to integrate pattern processing in the context of the PDP (parallel distributed processing) approach and symbol processing in the context of the conventional AI approach. Our model could also serve as a minimal platform for such a line of research, since the process in the simple brains of the artificial organisms must surely be transformation between stimulus patterns and the primitive words.

References

[1] K. Nakano, T. Ohmori, T. Arita and S. Takeda, "On a Self-Organizing System Obtaining the Ability to Exchange Information between Information Processors", SICE (Society of Instrument and control Engineers) 2nd Knowledge Engineering Symposium, 1984 (in Japanese).

[2] K. Nakano, Y. Sakaguchi, R. Isotani and T. Ohmori, "Self-Organizing System Obtaining Communication Ability: Primitive Model for Language Generation", Biological Cybernetics, Vol. 58, pp. 417-425, 1988.

[3] C. G. Langton, "Editor's Introduction", Artificial Life, Vol. 1, No. 1/2, pp. v-viii, The MIT Press, 1994.

[4] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, 1989.

[5] K. Nakano, "Associatron - A Model of Associative Memory", IEEE Trans. Syst., Man & Cybern. SMC-2, 3, pp. 381-388, 1972.

[6] D. O. Hebb, "The Organization of Behavior", Wiley, New York, 1949.

[7] B. MacLennan, "Synthetic Ethology: An Approach to the Study of Communication", Artificial Life II, pp. 631-658, Addison Wesley, 1991.

[8] G. M. Werner and M. G. Dyer, "Evolution of Communication in Artificial Organisms", Artificial Life II, pp. 659-687, Addison Wesley, 1991.

[9] D. H. Ackley and M. L. Littman, "Atrium in the Evolution of Communication", Artificial Life IV, pp. 40-48, The MIT Press, 1994.

[10] T. Arita, K. Unno and K. Kawaguchi, "A Model for Evolution of Communication in a Population of Artificial Organisms", forthcoming in Transactions of IPSJ (Information Processing Society of Japan), 1995 (in Japanese).

[11] K. Nakano, "Associatron: A Model of Associative Memory and its Intelligent Information Processing", Shoko-Do, 1979 (in Japanese).