SciencePG
Science Publishing Group

Each Role of Short-term and Long-term Memory in Neural Networks

Seisuke Yanagawa

OptID, Machida, Tokyo, Japan

Email address:
s_yanagawa@jcom.home.ne.jp

Abstract: Based on known functions of neuroscience the neural network that performs serial parallel conversion and its inverse transformation is presented. By hierarchy connecting the neural networks, the upper neural network that can process general time sequence data is constructed. The activity of the upper neural networks changes in response to the context structure inherent in the time series data and have both function of accepting and generating of general time series data. Eating behavior in animals in the early stages of evolution is also processing time series data, and it is possible to predict behavior although be limited short term by learning the contextual structure inherent in time series data. This function is the behavior of so-called short-term memory. Transition of the activation portion in this type of operation is illustrated. Although status of nervous system of the animal change according to the recognition by sensory organ and to the manipulation of the object by muscle in the vicinity of the animal itself, the evolved animals have in addition another nervous system so-called long-term memory or episodic memory being involved experience and prediction. The nervous system of long-term memory behaves freely but keeping consistency of the change in the environment. By the workings of long-term memory, lot of information are exchanged between fellows, and lot of time series data are conserved by characters in human society. In this paper, the model of the transfer of data between different nervous systems is shown using the concept of category theory.

Keywords: Short-term Memory, Long-term Memory, Serial Parallel Conversion, Parallel Serial Conversion, Mirror Neuron, Prediction, Category Theory

1. Introduction

If you look back at the human history of engineering progress, you will find that there are several stages. The use of wheels and gears, the creation of new energy from steam engines to nuclear power, and the changing information transmission methods from flags or smoke to the internet have been the catalyst for the beginning of each stage. In parallel with advances in technology, metals, chemicals, and even semiconductors have emerged as a new material for new products.

It is human intelligence that gave birth to the technology. But the structure of our brain has not changed since tens of thousands of years ago. Furthermore retroactively, although there is a quantitative difference in each part between brain of ape who does not speak the language and our brain, but our brain is consisted of same material. In addition, modern people have not advanced by having an element in the brain

that performs floating-point arithmetic at high speed. In other words, the nervous system of animals called advanced higher animals is locally same as very primitive animal's nervous system. Advance of the amount and complexity of the connection on nervous system must have been enabled to get more advanced processing. In this paper, a new neural network and its behavior are presented based on the above idea.

In Chapter 2, the circuits by combining the basic functions of neural circuits for serial parallel conversion and inverse of the conversion are shown. Although for learning process Hebb rules is used on the circuits, operations such as back propagation and Markov process are not used. By hierarchical connection of the circuits, it is possible to accept and generate general time series data. As an example, a neural network that can accept and generate multi step time series data such as indispensable for eating behavior is presented.

Chapter 3 describes the transition of the activated parts of the neural network corresponding the changes of the environment, the short-term memory and the long-term memory. The nervous system related to short-term memory is activated in synchronization with environmental events, but the nervous system involved in long-term memory is highly layered to form an image corresponding to past and future events. Acceptance and generation of time series data of the nervous system of long-term memory is carried out in consistent with short-term memory. In this paper, the model of the transfer of data between different nervous systems is shown using the concept of category theory.

2. Basic CIRCUIT That can Be Extended Deductively

Even bacteria like animals in the early stages of evolution must have some eating behavior such as moving relying on light and smell to search for food and determine whether they can be eaten. Whether the eating behavior is evaluated as intellectual behavior aside, it is possible to mimic this degree of behavior by electronic work at the junior high school level by combining sensors and logic ICs. The number of logical elements used may not differ much from the sum of nerve cells in insects or zooplankton. In the following, rather than the engineering usefulness, to extend the function of the circuit according the evolution of the neural system of the animal. The eating behavior of animals evolved from the aforementioned animals is composed of time series of actions such as extending arms toward the target, opening the palm of hand when approaching the target, closing the palm to grasp the target and bringing to the mouth. The same is true for the recognition process. We recognize five coins as the sum of two sets and three sets when they are shown. The same applies to general figures. Many East Asian recognize kanji as a combination of parts. The recognition technology of the figure has evolved by the neural network which starts with the perceptron, but it is forced to judge by the relation of the part and the whole as in the example of kanji when the scale of the figure increases. In other words, a trade-off is made between the recognition power of the part in the complex figure recognition and the processing power of the data sequence of the recognition result, and by learning the animal might got the most efficient processing method. On the development of neural networks that deal with general time series data, it is important to proceed without losing affinity with knowledge of neuroscience. It is desirable that the behavior of any neural network system can be expressed by combinations of the simple action parts of animals in the early stages of evolution. Deductive logical development is desired. In this chapter, it is presented that arbitrary time series data can be divide into basic sequences as the logical basis of neural networks. Next, by providing a two-way function to the neural network both serial parallel conversion and vice versa on basic sequences is realized. Finally, it shows the hierarchically connected neural network that can process for general time series data.

2.1. Dividing Time Series Data

Any time series data consisted finite type of element can be divided into multiple subsequences where the same element does not appear more than once. The divided subsequence is defined the basic subsequence. The neural circuit corresponding to each basic subsequence can be easily configured in the neural network, leading to the realization of processing by hardware of general time series data.

For simplicity, assume time series data consisted of 10 type elements from a_0 to a_9. Example shown in Figure 1 shows that the time series data (1) arranged by randomly selecting the elements is divided into five basic subsequences (2).

$$a_1 a_7 a_4 a_6 a_6 a_0 a_6 a_5 a_1 a_3 a_7 a_8 a_9 a_9 a_9 a_7 a_5 a_4 a_1 \qquad (1)$$

$$\downarrow$$

$$|a_1 a_7 a_4 a_6| \; |a_6 a_0| \; |a_6 a_5 a_1 a_3 a_7 a_8 a_9| \; |a_9| \; |a_9 a_7 a_5 a_4 a_1| \qquad (2)$$

Figure 1. Any time-series data has hierarchic structure.

The dividing is done by the following procedure.
(1) The first element is the beginning of the first subsequence. In given example, the leading element is a1, followed by a7, a4, a6 and a6. When the following conditions are true, the element allocate to the top of new subsequence.
(2) If the same element exists in already divided subsequence. In this example a6 is the concerned element.
(3) If the maximum length of subsequence is defined, a new subsequence is allocated after the subsequence that reaches the maximum length, add new element to the concerned new subsequence. The subsequences divided by above procedure are defined as the basic subsequence.

Figure 2 shows the affinity with the neural circuit. It is a neural network having an input consisting of a plurality of bits are shown. When the first data c_0 is received activate the bottom. For the next data c_1 additionally activates the elements which has been activated by the first data c_0. Because the elements activated by c_1 is randomly connected to input, not all elements activated by c_0 are additionally activated. Four portions are activated in the Figure 2. Other portion will return to the initial state because no activation factors (may be activated by another time series data).
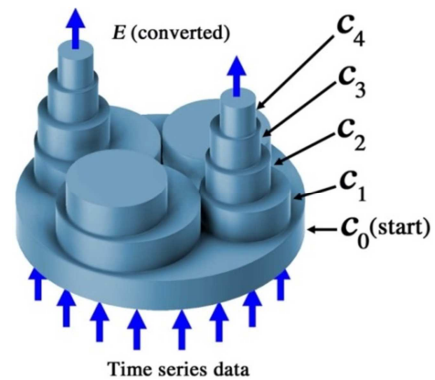

Figure 2. Changes of activation portions in the neural network when time series data is received.

By receiving c_2, c_3 and so on the activated portions become narrow and narrow. The output of the element holding the activity when receiving the last c_4 of the time series data is the recognition result of the time series data $c_0c_1c_2c_3c_4$. It may be seen as the output of the connected AND logic element. The number and their position of the activated elements is the conversion output corresponding to the serial input [1].

2.2. Bidirectional Conversion

In the neural network shown in Figure 2, the elements are activated one after another by the time series data (in this case, the basic subsequence) from below, and the result is output upward. Its output is the result of the serial parallel conversion, it is the result of AND logic of the output of the activated elements. The elements involved in the conversion are still activated at the time of output. Therefore, the couplings between the elements are enhanced (Hebb rule) by repeating this conversion, as a result the elements involved in the conversion will be activated by only receiving the first element of the time series data. This operation is a generation of (learned) time series data. that can be said a conversion of parallel to serial triggered by the first data. On the flow direction of data, this parallel to serial conversion is upside down with the serial to parallel conversion described above, but the basis of neural network operation is the same. When the state transition diagram of the Figure 3 is seen as a serial parallel conversion, the bottom is inputs connected to such sensory organs. After the first data reception, the connected elements are activated as described above. However, the conversion does not complete unless the following data are not coming. At each stage of the transformation, there is a data wait state, and the state returns to sleep state after a time limit.
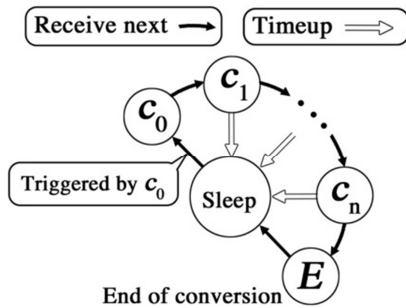

Figure 3. *State transition diagram of conversion.*

On the other hand, when the state transition diagram of the Figure 3 is seen as a parallel serial conversion, another waiting state is needed. In the spontaneous operation (here by voluntary muscle) performed reactions such as changes of weight feelings from muscle or joint are accompanied by. That is, the operation with no reaction within waiting time is invalid and aborted. The essence of the neural network which performs the parallel serial conversion and the neural network which performs the serial parallel conversion is the same except that the conversion result comes out downward or upward. When all data output is finished, a completion signal

is issued upward. On other words, both operations of the serial parallel conversion and the reverse conversion are essentially the same on the point both operations are triggered by the first data and proceed waiting for input state change. Both state transition diagram is shown in the existing Figure 3. For each basic subsequence a neural network that accepts and generates the basic subsequence is considered. This neural network is called a basic unit.

2.3. Hierarchical Connection of Basic Units

Since general time series data is consisted of hierarchy of basic subsequences, basic units can process general time series data by identifying the outputs of lower layer basic units as the new time series data. Animal's behavior is considered a time series data consisted of the data that is couple of behavior data and received data. Received data comes from the environment that encourages some judgment. The example shown at the beginning of this chapter is also a time series data that is consisted of extending arms, opening the palm, closing the palm to grasp and so on. In the Figure 4 activity changes of each unit on processing the time series data of eating behavior is shown.
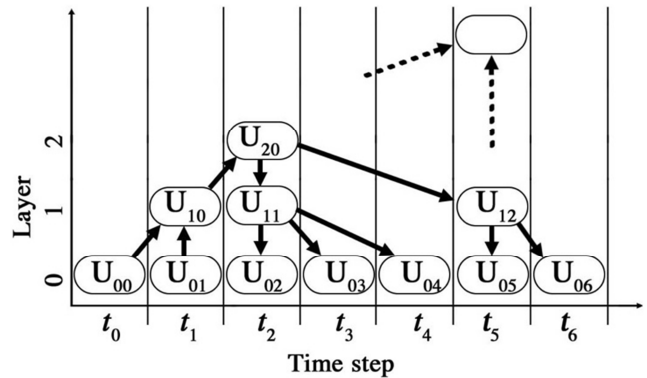

Figure 4. *Activity changes of each unit at eating.*

The time series data handled in the first 2 steps are consists of various stimuli of the internal and endocrine system, as well as data that captured by sensory organ from the environment in which the animal is placed. If the state of the unit becomes active (hungry), it has been shown to activate one after another the lower layer units that achieves eating behavior in the rest steps. After time sequence data learning, animals are possible to quickly start the operation that continues only by receiving the beginning of the time series data to be corresponding a slight sign. It can be said as "stance to the event" that most animals have [2].

Even during sleep, multiple activated areas move over the tree of the context of state change in the brain. As neural circuits evolve, the movement of each element is captured and accumulated as a context, and it becomes possible to adapt to changes in the new environment. In that case the movement of the element might be made in the highly layered nervous system. And in the nervous system, the eating behavior shown above will be nothing more than a brief occurrence that appears in a constantly continuing life. The reason why drawn

a hierarchy further in the upper part of the Figure 4, is to suggest existence of the nervous system that processes intellectual judgment. The movement will be mentioned in the next chapter. Neuroscientist Damasio calls "image" the internal representation built in the nervous system by stimulation inside and outside the body. And argues that an evolved animal has at the heart of the nervous system an elaborate network, which is the brain [3].

In brief, images were advantageous even if an organism were not conscious of the images formed within it. The organism would not yet be capable of subjectivity and would be unable to inspect the images in its own mind, execution of a movement; the movement would be more precise in terms of its target and succeed rather than fail.

3. Environment, Long-term Memory and Short-term Memory

Animals in the early stages of evolution will spend most of their living time obtaining food and avoiding danger. Animals with some evolved sensory organs must recognize, for example, environmental changes from sunrise to sunset as repetitions of a time series data. However, recognition is limited to the area of the vicinity that the sensory organ catches. Before long, animals evolve to be able to act in groups. They exchange information with their peers through squeals and gestures to enhance their ability to survive as a species.

In order to be able to exchange information, it is necessary to transmit and reproduce the act. The ability to imitate fellow's action is indispensable. Imitating is first step of learning the behavior of the fellows who is transmitting information. The nervous system which is involved in the imitating function is called mirror neurons. Imitating is the basis of fellow empathy and group behavior, and will develop to ethical emotions such as mercy and encouragement [4]. Human beings have adopted language as a means of exchanging information between peers, and began to exchange vast amounts of time series data. The object to be drawn is the change of all things, the joys and sorrows of livings, and their hope for the coming future. In addition, it spreads to the story beyond space and time. One of the factors that made the language possible to dramatically expand the range of expression is that it can be used as expression of object even if it is not near. For example, a family might have conversation like this toward you who is about to have a birthday the next day. "Tomorrow I'd like to buy a birthday cake at Store A to celebrate you with everyone." "Let's ask store A to put a chocolate plate on the cake with your name on it."

And you'll think of the shape of the cake and the action of lifting the chocolate plate and mouthing it. On your birthday day, you may want to see the cake in front of you, identify the cake from store A, which was the one that was the subject of a conversation with your family the day before, and worry about the difference from the previous day's expectations. Because

the episode about the cake is remembered, communication is possible and feelings are transmitted. If there is a problem in the episodic memory, it causes difficulties in social activities. Two types of time series data can be considered in the above situation. One is time series data based on visual information of the cake in front of you and muscle movements that manipulate chocolate plates, which is produced by the nervous system that animals have from the early stages of evolution. The other is generated by the nervous system that is called "image" by Damasio, is time series data based on the shape of cake produced from family conversations the day before. While the behavior of the former nervous system is reflection of the visual information and movement of the objects nearby, the latter nervous system not only no needs to be synchronized with the former, but also moves independently. Like talking about your childhood while eating cake. However, if the nervous system of the episodic memory is activated by remembering the A shop and the chocolate plate while looking at the cake, the difference between reality and the expectation might become a problem. In order to avoid confusion, the episode must be corrected by reality.
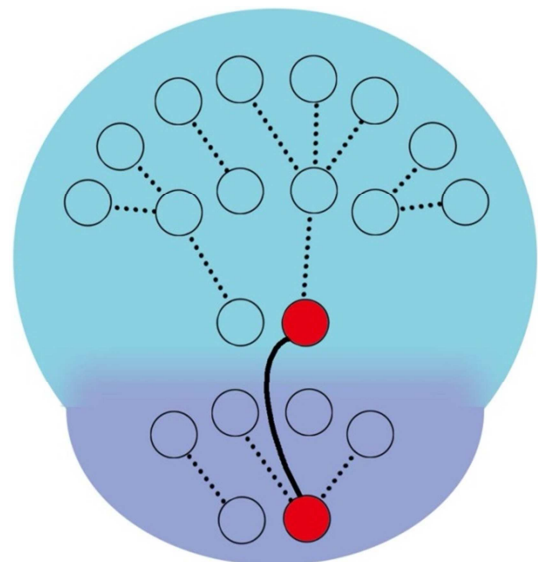

Figure 5. *Interaction between two nervous system.*

Figure 5 is an illustration for showing the state change of each part in the neural network. The upper part shows the part related to episodic memory, and the lower part shows the part related to the short-term memory. Stimulus from the sensory organs from the bottom becomes time series data and is transmitted to the upper part. The part shown in red is a part that is particularly activated, the lower red disk is the part activated by the visual data of the chocolate plate that placed in front, the upper red disk is the part that is activating by recalled episode about the chocolate plate. The connection between the two sides is enhanced (Hebb rule) as indicated by a bold line, and visual information is copied to the top to reinforce the episodic memory. As a result, the shape of the chocolate in the episodic memory is identified as the thing seen in front.

When studying objects that are intertwined with portion and whole system such as the nervous system, the idea of category theory can be incorporated to develop the whole without focusing on the details of the object. The following is a description of Figure 6 associated with the definition of the category shown in Tom Leinster "Basic Category Theory" [4].

A category N consists of:

[Def.1] a collection ob (N) of objects;

In this paper, object corresponds to nerve cells in

neuroscience and in neural networks corresponds to basic units (or combinations thereof) described in Figure 6.

[Def.2] for each $A, B \in$ ob (N), a collection $N (A, B)$ of morphism from A to B;
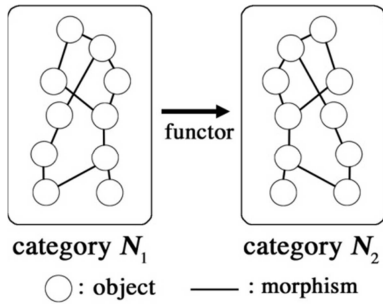

Figure 6. *Illustration of category and functor in neural networks.*

In neuroscience, the morphism corresponds to synapses and axons, and is responsible for the transmission of information between objects. In neural networks, it is a coupling element with the property that the binding is enhanced as the activity of both ends increases.

[Def.3] for each $A, B, C \in$ ob (N), a function

$$N (B, C) \times N (A, B) \rightarrow N (A, C),$$

called composition;

[Def.4] for each $A \in$ ob (N), an element 1_A of $N (A, A)$,

called the identity on N, satisfying the following axioms:

[Ax.1] associativity: for each $f \in N (A, B)$, $g \in N (B, C)$ and

$h \in N (C, D)$, we have $(h \circ g) \circ f = h \circ (g \circ f)$;

[Ax.2] identity laws: for each $f \in N (A, B)$, we have

$$f \circ 1_A = f = 1_B \circ f.$$

The composite of the coupling is considered to be a hierarchical connection of the objects, and the identity, which is considered a special morphism, can correspond long axons extending beyond hierarchies.

Next, we introduce a function called functor between different categories. The object and morphism that make up the source category are connected by a function to the object and morphism of the target category. According to the Hebb law, if there are two nervous system activated, the binding between the elements in the two categories will be enhanced. Therefore, the existence of the above functor would be considered naturally. Thus, the behavior of both categories can be migrated to each other. The process of migration can be described mathematically using free functor or free

construction functor.

4. Conclusion

It has been vaguely thought that the nervous system responsible for long-term memory may be in a different place from the nervous system responsible for short-term memory. In this paper, short-term memory and long-term memory are both regarded as time series data brought about by activity in the brain. It is a short-term memory to be involved in the recognition and manipulation of the thing in the vicinity of the animal itself. And the time series data brought by other brain with highly hierarchically connected structure   activities is concerning long-term memory. It is necessary that the contents of both memories are consistent in the animal's life. Consistency is required when the objects in which both memories are involved are the same. The process of taking the consistency between the two memories was explained using the idea of a free functor in category theory. Of course, the description does not convey all the features of the hippocampus. In the study of memory and language, it is a meaningful theme that covers the homogeneity of behavior between parts of the nervous system [6, 7].

From an engineering point of view, the basic unit can be realized by a small microprocessor, and the neural network consists of multiple randomly connected units. Only units contained in the tree structure corresponding to the context of the behavior can work. Each unit has the same function, but the connections between the units are not the same. From among the random connections, the necessary connections for the desired operation are selected and enhanced, and the target function is realized. It can be said that the essence of the logic of the operation exists not in the basic unit, but in the connecting situation among basic units selected from randomly initialized connection. Therefore, even if the circuit is partially damaged, it is possible to supplement by learning the function of the peripheral circuit is lost. This process is close to the rehabilitation process of the brain that has had a stroke. There may be cases of errors in the accuracy of the operation compared to the circuit using the existing logic IC because there is a probabilistic part, but the bud of a new strategy might be hidden in the vicinity of the malfunction. The bud is also felt in the judgment and the movement in our daily life.

From the viewpoint of neuroscience, even if there are no "parts" equivalent to the basic unit in the process of stimulation from the sensory organs propagating to the cerebral cortex, axons parallel to the propagation direction connect between the layers, and it can be said that it is a passage of serial parallel mutually converted information. On the other hand, it is considered that the axon perpendicular to the propagation direction is a connection such as causing self-oscillation related to the generation of serial data.

Constructing neural networks based on known functions of neural circuits, including the Hebb rule, and considering the correspondence between the movements of biological neural circuits will lead to new discoveries in both fields.

References

[1] S. Yanagawa, "Neural Network That Learns Sequential Processing and Predicts by the Context." In The 28th Annual Conference of the Japanese Neural Network Society (JNNS2018).

[2] S. Yanagawa "From Bacteria, a Consideration of the Evolution of Neural Network" Journal of Mechanics Engineering and Automation 1 (2019) p. 17-23.

[3] A. Damasio The Strange Order of Things Pantheon Books, New York, 2018 p. 62.

[4] T M. Iacoboni, "Mirroring People." Picador 33. 2008, Chapter 4, p 106.

[5] T. Leinster "Basic Category Theory" Cambridge University Press 2017 p. 10.

[6] D. C. Dennet, "From Bacteria to Bach and Back: The Evolution of Minds." In Penguin Books, 2018.

[7] S. Yanagawa "A neural network that processes time series data in which element data is linked to symbols Third International Workshop on Symbolic-Neural Learning (SNL-2019) Poster presentation.

[8] S. Yanagawa. 2017, "Learning Machine That Uses Context Structure to Search Policy." https://jsai.ixsq.nii.ac.jp/.SIG-AGI-007-07.

[9] S. Dehaene, Consciousness and Brain Penguin Books, 2014.

[10] W. J. Freeman "How Brains Make up Their Minds" Weidenfeld & Nicolson Ltd. 1999.

[11] Hoon Keng Poon, Wun-She Yap, Yee-Kai Tee, Wai-Kong Lee, Bok-Min Goi "Hierarchical gated recurrent neural network with adversarial and virtual adversarial training on text classification", Neural Networks 119 (2019) 299-312.

[12] L. Andrew Coward, Ron Sun, "Hierarchical approaches to understanding consciousness", Neural Networks 20 (2007) 947-954.

[13] Qianli Ma, Wanqing Zhuang, Lifeng Shen, Garrison W. Cottrell, "Time series classification with Echo Memory Networks", Neural Networks 117 (2019) 225-239.

[14] Paolo Arena, Marco Cali, Luca Patane, Agnese Portera, Roland StraussSC, "Modelling the insect Mushroom Bodies: Application to Sequence learning", Neural Networks 67 (2015) 37-53.

[15] Michael j. Healy, Thomas P. Caudill, "Episodic memory: A hierarchy of spatiotemporal concepts", Neural Networks 120 (2019) 40-57

[16] Kenji Doyaa & Tadahiro Taniguchi, "Toward Evolutionary and Developmental Intelligence", Current Opinion in Behavioral Sciences Volume 29, October 2019, Pages 91-96.

[17] Ann Sizemore, Chad Giusti, Ari Kahn, Richard F. Betzel, Danielle S. Bassett "Cliques and Cavities in the Human Connectome", Cornell University arxiv.org/abs/1608.03520, 2016.

[18] Hiroshi Yamakawa, "Attentional Reinforcement Learning in the Brain", New Generation Computing, January 2020.