

Dual-Network Memory Model for Temporal Sequences

Motonobu Hattori, Rina Suzuki

Abstract—In neural networks, when new patterns are learned by a network, they radically interfere with previously stored patterns. This drawback is called catastrophic forgetting. We have already proposed a biologically inspired dual-network memory model which can much reduce this forgetting for static patterns. In this model, information is first stored in the hippocampal network, and thereafter, it is transferred to the neocortical network using pseudopatterns. Because temporal sequence learning is more important than static pattern learning in the real world, in this study, we improve our conventional dual-network memory model so that it can deal with temporal sequences without catastrophic forgetting. The computer simulation results show the effectiveness of the proposed dual-network memory model.

Keywords—Catastrophic forgetting, dual-network, temporal sequences.

I. INTRODUCTION

WHEN a neural network is trained on one set of patterns and then when it later attempts to add new patterns to its repertoire, catastrophic interference or complete loss of all of its previously learned information may result. This type of radical forgetting is unacceptable for both a model of human memory and practical engineering applications. One of the typical solutions to catastrophic forgetting is *interleaved learning*, that is, mixing in patterns of the previous set during training on a new pattern set. However, this solution requires the unrealistic assumption of permanent access to all patterns on which the network was previously trained on. Thus, to reduce catastrophic forgetting without straight rehearsal, several methods have been proposed. Most learning techniques to reduce catastrophic forgetting have relied on reducing the overlap of internal representations [1], [2], [3], [4].

On the other hand, various authors have developed models that employ pseudopatterns rather than the real patterns that were learned previously [5], [6], [7], [8]. Among them, we have recently proposed a biologically inspired dual-network memory model [9]. This model is composed of two distinct networks: a hippocampal network for early processing and a neocortical network for long-term storage. This model employs a chaotic neural network [10] in CA3 of the hippocampal network, and the information stored by the hippocampal network is transferred to the neocortical network by chaotic recall of the hippocampal network. Because previously learned original patterns can be extracted with chaotic recall, we have shown that our dual-network model can significantly reduce catastrophic forgetting [9].

M. Hattori is with Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, JAPAN. (E-mail: m-hattori@yamanashi.ac.jp, phone & fax: +81-55-220-8766).

R. Suzuki is with NTT DOCOMO, INC., Tokyo, JAPAN. (E-mail: rina.suzuki.wb@nttdocomo.com).

Although it is believed that the hippocampus has an important role for processing not only semantic memory but also episodic memory in the human brain [11], our previous model can deal with only static patterns. Because much of what we do consists of mastering the serial order of pattern sequences: from calling phone to starting a car. Therefore, it is very important to be able to learn serially many temporal sequences of patterns without catastrophic forgetting. Hence, in this paper, we extend our previous model to deal with temporal sequences of patterns without catastrophic forgetting, and compare its performance with that of the conventional model [12].

The rest of this paper is organized as follows. In Section II, we briefly review the conventional dual-network memory model for temporal sequences [12]. Section III explains the proposed dual-network memory model. In Section IV, we show the experimental results. Finally, conclusions are given in Section 5.

II. CONVENTIONAL DUAL-NETWORK MEMORY MODEL

Fig. 1 shows a reverberating simple recurrent network (RSRN) which is designed to learn temporally ordered sequences of patterns $S(1), S(2), \dots, S(t), \dots, S(n)$, where t is the time parameter [12]. A RSRN involves adding *associative nodes* to the output layer of a simple recurrent network, or Elman network [13]. The input to a RSRN consists of the sequence item $S(t)$, and the context input $H(t-1)$.

The conventional dual-network model for learning temporal sequences consists of two identical RSRNs, NET 1 and NET 2 [12]. New sequences are learned only by NET 1, while NET 2 stores the previously learned information. The

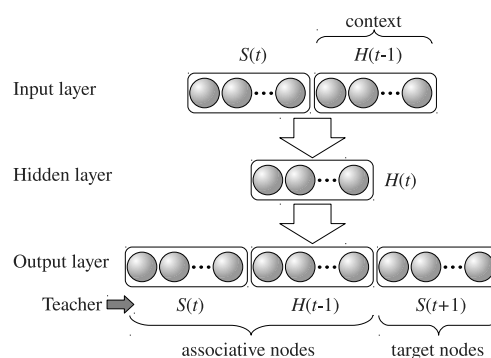


Fig. 1. Structure of the conventional dual-network memory model for temporal sequences [12].

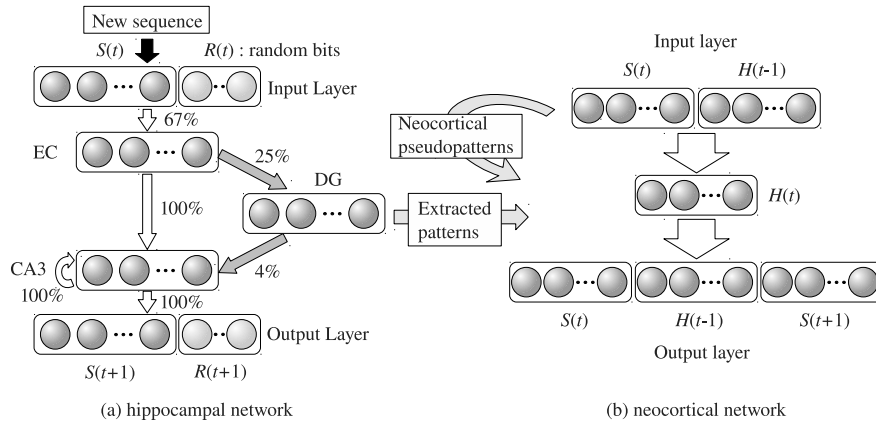


Fig. 2. Structure of the proposed dual-network memory model for temporal sequences. In the hippocampal network, the connections shown in gray arrows are used only in learning, and the connectivity between the two regions is shown in percentage.

learning procedure of the RSRN dual-network is as follows. Suppose that some previously learned sequences are already stored in NET 2. When a sequence, $S = S(1), S(2), \dots, S(n)$, is presented to NET 1, the network makes a single pass through the entire sequence, updating its weights once for each item in the sequence. This defines one learning *epoch*. In the mean time, a random input i is presented to the input layer of NET 2 and fed through the network to the output layer. Only the activation values of the associative nodes in the output layer are used to constitute a new input, i' . This vector i' is then given to the network in order to produce a pattern of activation on the autoassociative output nodes, i'' . After a number R of these *reverberating* cycles through the network, the final *reverberated* input i^R is sent through the network, and the activation vector on all output nodes, o , is used to produce a pseudopattern, $i^R \rightarrow o$. This pseudopattern reflects the contents of the previously learned information stored in NET 2. NET 2 generates M_1 pseudopatterns. For each NET 2 pseudopattern, NET 1 performs on feedforward backpropagation learning pass. Once this is completed, a new learning epoch starts and NET 1 makes another pass through the sequence S . NET 2 generates new M_1 pseudopatterns, each of which is learned by NET 1. This process continues until S is learned to criterion by NET 1. This is called the *interleaving phase*.

There is also a *transfer phase* during which the information learned in NET 1 is transmitted to NET 2. This involves NET 1 generating M_2 pseudopatterns and, for each of these NET 1 pseudopatterns, NET 2 performs a single learning pass, thereby transferring information stored in NET 1 to NET 2.

III. NOVEL DUAL-NETWORK MEMORY MODEL FOR TEMPORAL SEQUENCES

Fig. 2 shows the structure of the proposed dual-network memory model. We have modified the conventional dual-network model for static patterns proposed in ref.[9] so that it can deal with temporal sequences. The proposed model consists of two distinct neural networks: hippocampal and neocortical networks. Information is first stored in the

hippocampal network, and thereafter, it is transferred to the neocortical network. The fundamental difference between our conventional dual-network model for static patterns [9] and the proposed one is as follows.

- 1) Random-bit nodes are added to both Input and Output layers in the hippocampal network. For each pattern of a sequence, $S(t)$, a randomly generated pattern, $R(t)$ is concatenated in order to learn an ambiguous sequence (e.g. one that contains two identical patterns).
- 2) In the conventional model [9], chaotic neurons are employed in CA3 layer in the hippocampal network in order to recall stored patterns dynamically. However, we replace them with normal (non-chaotic) neurons because a series of patterns should be recalled in order in the dual-network model for temporal sequences.
- 3) In order to store temporal sequences in the neocortical network, we substitute a RSRN for a feedforward network used in the conventional model.

A. Learning for Temporal Sequences

Suppose that $S = S(1), S(2), \dots, S(n)$ shows a sequence to be learned. In the learning of the proposed dual-network memory model, a random extra pattern, $S(n+1)$ is added to S . For each pattern, $S(t)$, a randomly generated pattern, $R(t)$ is concatenated. Then, as shown in Fig. 2(a), two consecutive patterns are learned by the hippocampal network, where $S(n+2) = S(0)$.

Although the connection weights between Input and EC are fixed at 1, other weights are learned as follows.

EC–DG, EC–CA3, DG–CA3:

We have employed Oja's rule [14] because it constrains the divergence of weights:

$$w_{ij}(t+1) = w_{ij}(t) + \eta y_j (x_i - y_j w_{ij}(t)), \quad (1)$$

where x_i is the output of a neuron in a layer, y_j is the neuron in the subsequent layer, w_{ij} denotes the connection weight between these neurons, and η is the learning rate.

CA3:

Weights are learned using the Hebbian learning:

$$w_{ij}(t+1) = w_{ij} + x_i x_j, \quad (2)$$

where $w_{ij} = w_{ji}$ and $w_{ii} = 0$.

CA3-Output:

Weights are learned using the Hebbian learning:

$$w_{ij}(t+1) = w_{ij}(t) + x_i y_j, \quad (3)$$

where x_i and y_j show the output of the i th neuron in CA3 and that of the j th neuron in Output, respectively.

Because each learning rule in Eqs.(1)–(3) is based on Hebbian learning, the proposed hippocampal network acquires a new sequence much more rapidly in comparison with those learned by the backpropagation algorithm [12]. Moreover, Hebbian learning does not suffer from catastrophic forgetting when the number of patterns to be stored is small. Thus, catastrophic forgetting can be much reduced in the proposed hippocampal network and it functions as a short-term memory.

A stored sequence in the hippocampal network is transferred to the neocortical network as follows. Suppose that some previously learned sequences are already stored in the neocortical network.

- 1) A random input is applied to Input layer in the hippocampal network.
- 2) On the basis of the evidence from electrophysiological experiment [15], only the path Input→EC→CA3→Output is used in recall. Neurons in CA3 layer change their states asynchronously until the state of the layer reaches a stable state. Then, the output of the network is obtained.
- 3) The output is applied to Input layer.
- 4) Repeat 2) and 3) until the whole sequence is recalled.
- 5) Put patterns in order by using a random extra pattern, $S(n+1)$.
- 6) Remove $S(n+1)$ from the recalled sequence and random bits from each pattern.
- 7) The neocortical network generates M_1 pseudopatterns. The resultant sequence in 6) is learned by the neocortical network with these pseudopatterns.

In the proposed dual-network memory model, the original sequence learned by the hippocampal network may be available for the learning of the neocortical network. This can much contribute to reduce catastrophic forgetting.

IV. COMPUTER SIMULATION RESULTS

In computer simulation, we set the dimension of a pattern in a sequence at 100. Each pattern in a sequence consists of 0s and 1s with a selection probability of 0.5 except the third experiment. The number of random-bit nodes in the proposed hippocampal network was set to 50, and the number of hidden neurons in the RSRN was set to 50. The numbers of pseudopatterns, M_1 and M_2 were set to 50 and 5000, respectively.

To evaluate a network's ability to correctly reproduce the appropriate outputs, a performance measure, called *goodness* [5],

[7], [9], [16], was adopted. Let o_i be the output value of the i th output neuron of NET 2 or the neocortical network, and let t_i be the corresponding component of the desired pattern. Then, the goodness g is defined as follows:

$$g = \frac{1}{N} \sum_{i=1}^N (2o_i - 1)(2t_i - 1) \quad (4)$$

where N is the number of target nodes in Output layer in RSRN. A goodness value 1 indicates a perfect match between the calculated and desired outputs, whereas 0 indicates chance level.

The results of computer simulation in this section are based on 20 trials.

A. Catastrophic Forgetting in Complex Sequences

In this experiment, we used complex sequences called *SOCs: second-order conditional sequences* [12], in which no single element in the sequence can predict its successor. The complexity of these sequences derived from their structure consisting of four distinct items appearing in an order such that every item is immediately followed by one of the three other possible items with equal frequency. A single item cannot by itself predict the following item better than chance. In other words, two consecutive items are always required to predict the next one. We created the following two SOC sequences:

$$S_1 : A \rightarrow B \rightarrow C \rightarrow A \rightarrow D \rightarrow C \rightarrow D \rightarrow B \rightarrow A \rightarrow C \rightarrow B \rightarrow D \rightarrow A \quad (5)$$

$$S_2 : E \rightarrow F \rightarrow G \rightarrow E \rightarrow H \rightarrow G \rightarrow H \rightarrow F \rightarrow E \rightarrow G \rightarrow F \rightarrow H \rightarrow E \quad (6)$$

where each alphabet shows 100 dimensional random pattern, and the length of each sequence is 13.

Fig. 3 shows mean goodness of each item of S_1 after S_1 and S_2 were learned in this order. In this experiment, only the first item of S_1 is given as an input. Therefore, the item in position 1 of S_1 is not shown in the figure because it is used to initialize recall. As shown in Fig. 3, the performance of the conventional dual-network is getting worse and worse

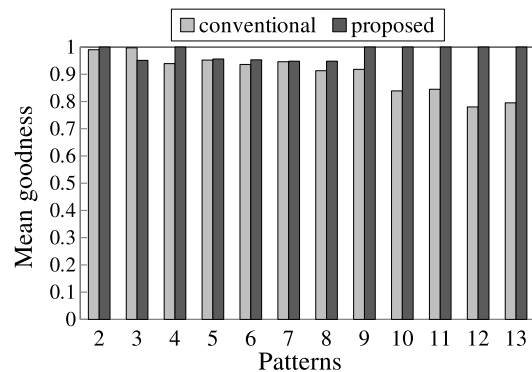


Fig. 3. Mean goodness for the previously learned SOC, S_1 after S_1 and S_2 were learned serially.

for items in the latter position of the sequence. In contrast, the proposed dual-network model shows much better performance in reducing catastrophic forgetting even for the second half of the sequence.

B. Catastrophic Forgetting in Multiple Sequences

In this experiment, five sequences, from S_1 to S_5 were sequentially learned by both the conventional and the proposed dual-network models, and we examined how both models could reduce catastrophic forgetting for old information. The number of items in each sequence was set to 10.

Fig. 4 shows mean goodness of items in each sequence. As shown in the figure, the proposed dual-network model significantly reduces catastrophic forgetting compared with the conventional model.

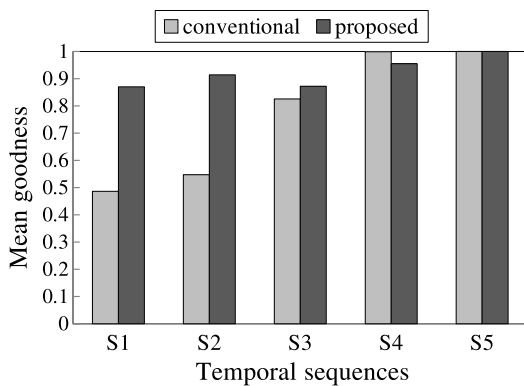


Fig. 4. Mean goodness for five sequences.

C. Catastrophic Forgetting in Similar Sequences

In this experiment, we examined the relation between the ability to reduce catastrophic forgetting and the similarity of sequence. We defined the similarity of sequence by averaged direction cosine as follows:

$$similarity = \frac{1}{P C_2} \sum_{m=1}^{P-1} \sum_{n=m+1}^P \frac{(S(m), S(n))}{\|S(m)\| \|S(n)\|} \quad (7)$$

where $S(m)$ shows the m th pattern of a sequence, P shows the number of patterns in a sequence, $(S(m), S(n))$ denotes the inner product of $S(m)$ and $S(n)$, and $\|S(m)\|$ denotes the norm of $S(m)$. In this experiment, we made two sets of sequences the similarities of which were 0.3 and 0.7 by changing the selection probability of 0s and 1s. Each set has three sequences with the same similarity, and these sequences were sequentially learned by both the conventional and the proposed dual-network models. Figs. 5 and 6 show the results of this experiment. In both cases, the proposed model shows much better performance in reducing catastrophic forgetting. As seen in these figures, the proposed model is much more insensitive to the similarity of sequence.

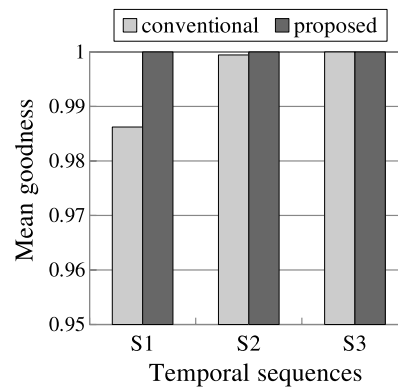


Fig. 5. Mean goodness for temporal sequences in which similarity was 0.3.

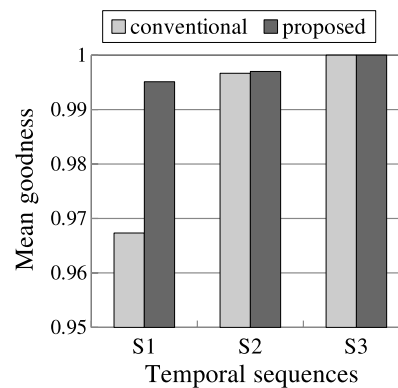


Fig. 6. Mean goodness for temporal sequences in which similarity was 0.7.

V. CONCLUSIONS

In this study, we proposed a novel dual-network memory model which can learn serially temporal sequences without catastrophic forgetting. In the conventional dual-network memory model [12], pseudopatterns are used in both the learning of NET 1 and that of NET 2. In contrast, in the proposed dual-network memory model, because the hippocampal network can extract a previously learned sequence, pseudopatterns are used only in the learning of the neocortical network. Therefore, the proposed model can much reduce catastrophic forgetting in comparison with the conventional one. The computer simulation results show that the proposed dual-network model has the following features:

- 1) It shows much better performance in reducing catastrophic forgetting for complex sequences.
- 2) It shows much better performance when the number of sequences to be stored is large.
- 3) It is much more insensitive to the similarity of sequence.

In future research, we plan to develop our model into a more biologically plausible one by using spiking neurons that should allow application of the model to practical problems.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 25330286.

REFERENCES

- [1] R. French, "Using semi-distributed representation to overcome catastrophic forgetting in connectionist network," in *Proceedings of the 13th Annual Cognitive Science Society Conference*, 1991, pp. 173–178.
- [2] —, "Dynamically constraining connectionist networks to produce distributed, orthogonal representations to reduce catastrophic interference," in *Proceedings of the 16th Annual Cognitive Science Society Conference*, 1994, pp. 335–340.
- [3] T. Yamada, M. Hattori, M. Morisawa, and H. Ito, "Sequential learning for associative memory using kohonen feature map," in *Proceedings of IEEE and INNS International Joint Conference on Neural Networks*, vol. 3, Washington, D.C., July 10–16 1999, pp. 1920–1923.
- [4] M. Hattori, H. Arisumi, and H. Ito, "Sequential learning for som associative memory with map reconstruction," in *Lecture Notes in Computer Science*, vol. 2130, Vienna, August 21–25 2001, pp. 477–484.
- [5] A. Robins, "Catastrophic forgetting, rehearsal, and pseudorehearsal," *Connection Science*, vol. 7, pp. 123–146, 1995.
- [6] R. French, "Pseudo-recurrent connectionist networks: An approach to the "sensitivity-stability" dilemma," *Connection Science*, vol. 9, no. 4, pp. 353–379, 1997.
- [7] B. Ans and S. Rousset, "Avoiding catastrophic forgetting by coupling two reverberating neural networks," *cademie des Sciences, Sciences de la vie*, vol. 320, pp. 989–997, 1997.
- [8] M. Hattori, "Dual-network memory model using a chaotic neural network," in *Proceedings of IEEE and INNS International Joint Conference on Neural Networks*, Barcelona, July 18–23 2010, pp. 1678–1682.
- [9] —, "Avoiding catastrophic forgetting by a biologically inspired dual-network memory model," in *Lecture Notes in Computer Science, Part II*, vol. 7664, Nov. 12–15 2012, pp. 392–400.
- [10] K. Aihara, T. Takabe, and M. Toyoda, "Chaotic neural networks," *Physics Letters A*, vol. 144, no. 6–7, pp. 333–340, 1990.
- [11] H. Eichenbaum, *The Cognitive Neuroscience of Memory*. New York: Oxford University Press, 2002.
- [12] B. Ans, S. Rousset, R. French, and S. Musca, "Self-refreshing memory in artificial neural networks: learning temporal sequences without catastrophic forgetting," *Connection Science*, vol. 16, no. 2, pp. 71–99, 2004.
- [13] J. Elman, "Finding structure in time," *Cognitive Science*, vol. 14, pp. 179–211, 1990.
- [14] J. Hertz, A. Krogh, and R. Palmer, *Introduction to The Theory of Neural Computation*. Redwood City: Addison-Wesley Publishing Company, 1991.
- [15] M. Yeckel and T. Berger, "Feedforward excitation of the hippocampus the trisynaptic pathway," in *Proceedings of the National Academy of Sciences of the USA*, vol. 87, 1990, pp. 5832–5836.
- [16] R. Ratcliff, "Connectionist models of recognition memory: constraints imposed by learning and forgetting functions," *Psychological Review*, vol. 87, no. 2, pp. 285–308, 1990.



Rina Suzuki received B.E. degree in Computer Science and Media Engineering from University of Yamanashi, Kofu, Japan, in 2013.

She is currently with NTT DOCOMO, INC., Tokyo, JAPAN.



Motonobu Hattori received B.E., M.E. and Ph.D degrees in Electrical Engineering from Keio University, Yokohama, Japan, in 1992, 1994 and 1997, respectively.

He is currently an Associate Professor with Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu, Japan, where he joined in 1997. From 2003 to 2004, he was a Visiting Researcher with the Center for the Neural Basis of Cognition (CNBC), Carnegie Mellon University, Pittsburgh, USA. His current

research interest is in neural networks, reinforcement learning and evolutionary computation.

Dr. Hattori is a member of the IEEE, JNNS, IEICE and IEE. He was a recipient of Niwa Memorial Award in 1996.