Future Prediction with Hierarchical Episodic Memories under Deterministic and Stochastic Environments

Yoshito Aota[1,2] and Yoshihiro Miyake[1]

[1] Tokyo Institute of Technology, Tokyo, Japan
miyake@dis.titech.ac.jp
[2] Yokohama National University, Yokohama, Japan
aota@ynu.ac.jp

Abstract. In agreement with Bond's suggestion, we consider that episodic memories are hierarchized autonomously by simple rule. In this research, our model solves maze tasks. Each episodic memory corresponds to the model's each track. In our previous research, we suggested that our model concatenates episodic memories into one long episodic memory. Our previous model showed successful prediction of any long periodical and deterministic environmental changes with editing (selecting and concatenating with adequate timing) stored episodic memories autonomously. However, the previous models could not select adequate actions under a stochastic environment like POMDPs. Here, we suggest hierarchical episodic memories implement into the model. It is shown that the model improved not only their action under POMDPs but also prediction of long-term environmental change and incremental learning.

Keywords: episodic memory, hierarchization, maze task, POMDPs, episodic future thinking, prediction, environment modeling, incremental learning.

1 Introduction

As Tulving [1] described, one of properties for episodic memory is temporal organization. From theoretical demand on a neural representation of episodic memories, Bond [2] claimed that since the number of events in one episodic memory is limited, episodic memories form hierarchies. Eichenbaum et al. [3] positioned rat's place cells as semantic memories and these are built from episodic experiences. Our former research demonstrated this process in a mathematical model [4]. Furthermore, it is well-known that activities of place cells are hierarchical according to the rat's experiences (e.g. [5][6][7]). Therefore, we agree with Bond's hierarchical property of episodic memories [2].

Rats can learn alternative maze task (e.g. [8]) of which the goal changes alternatively between a task A and a task B. We consider that hierarchical episodic memories should be constructed autonomously through experiences in order to reason the order AB. It is recommended to make a model creating hierarchical episodic memories aimed to predict future events that the model may encounter (about episodic future thinking and episodic memory, see [9]).

T. Huang et al. (Eds.): ICONIP 2012, Part I, LNCS 7663, pp. 251–259, 2012.

In our previous research [10][11], we suggested a model that makes and edits episodic memories. The unique point in our model was to concatenate stored episodic memories into one episodic memory with adequate timing and simple criteria, although many other researchers used episodic memories without editing (e.g. [12][13]. And also, Nuxoll and Laird [14] reviewed other researchers well and they suggested a model, which used simple episodic memories and showed possibilities of wide cognitive capabilities) or with editing but uses complex criteria (e.g. [15]). Our previous model could realize easy calculation by integrating some episodic memories into one episodic memory. The model predicted correct timing of task change like the order AB.

However, since integrated episodic memory spans longer time scale, it was not very useful for stochastic environments like Partially Observable Markov Decision Processes (POMDPs).

In this paper, we propose a model constructing hierarchy of episodic memories for deterministic and stochastic environments including POMDPs. We show the model can predict not only goal location in each maze task but also the order between tasks. Moreover, the model chooses reasonable actions in a stochastic environment by utilizing shorter episode, which is positioned lower in the hierarchy.

2 Maze Tasks

Fig. 1a) shows a cross maze task. The model can move forward, right, left, and backward. It takes one step. However, if there is a wall, the model does not choose its direction. We consider that convergence property would be the same even if the model takes wall into account, since the model encodes each episodic memory with one neuron that makes possible to calculate same time scale with short-term episodic memory as well as long term episode, though it will take more time until the model reach the goal. When the model reaches the end of branch, it moves to the start location S at the next step.

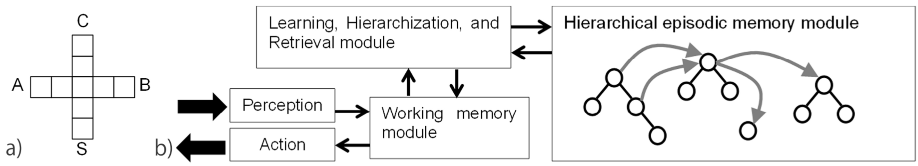

Fig. 1. a) Cross maze task. Goal is located in one of the route ends, A, B, and C. b) Basic structure of the model.

Although this maze is quite simple, it can realize three kinds of complexities, which the model would not know nor solve from the input itself.

One is “deterministic goal change” that the goal of each task changes deterministically and periodically. Second is “stochastic route change” that the west branch and the north branch in the cross maze exchange each other with some probability at the beginning of each task. Third is “incremental goal change” that the period of goal change becomes longer and longer according to the model’s learning progress.

Here, we do not consider optimal answer from start to goal, but evaluate whether the model could predict next goal location.

2.1 Deterministic Goal Change in the Maze Task

In order to confirm that our new model with hierarchical episodic memories shows the same capability as our previous model, which was without the hierarchical episodic memories [10][11], the new model must solve the same task as our previous ones.

The model must reach the goal located at the end of the route A as the first task (Fig. 1a). Then the location of the goal changes from A to B as the second task. The model must find the location of the new goal location which is on the route B as the third task. Like this, the goal location changes as task progresses. Moreover, this change shows periodic pattern, that is AABAABCABCAACAACBACB per one periodic pattern. The model is given such patterns many times. Finally, the model has to learn not only each task but also such task change pattern for correct prediction of goal locations.

In this paper, we call a periodic pattern of the goal change as “Task set”.

2.2 Stochastic Route Change

It would not make sense if the model tries to predict long-term changes, in case the stochastic environment changes but not the deterministic environment. In such a case, the model should shift from long-term prediction to short-term prediction. We have tested that the model can choose shorter and adequate actions under the POMDPs environment. We also assume that goal change is deterministic but branches in the maze change probabilistically. The model must extract the deterministic goal change from the stochastic environmental change.

In the task, change of the goal location shows a pattern AAB in Fig.1a. As a stochastic environmental change, we add probabilistic route exchange in this maze task. The routes A and C interchange by the probability of 20% at the beginning of each task. It is expected that, for instance, the model returns to the crossing intersection and chooses the correct route when the model predicts the model’s action would result in route A but actual route needs to be C.

2.3 Incremental Goal Change

In our previous works [10][11], we showed our model successfully transferred some parts of acquired task change patterns to a new pattern. Here, we confirm the same capability in our proposed model as well.

The model learns the goal change pattern ABC and ACB at the first and the second stage, then learns AABAABC and AACAACB at the third and the fourth stage, finally the model learns AABAABCABCAACAACBACB being the same as that of the task which is described in the section 2.1.

3 Model with Hierarchical Episodic Memories

Fig. 1b) depicts an abstracted view of the proposed model integrated with the hierarchical episodic memory architecture. The model first chooses a suitable episodic memory and keeps it in a working memory module (Fig. 1b). Next, the model performs actions directed by the episodic memory kept in the working memory module. When the model encounters an unexpected state, the model makes a new episodic memory in the working memory module and stores it in the hierarchical episodic memory module if it is new episodic memory for the model. Moreover, at that time, the model combines two or more episodic memories into one episodic memory if needed. They are concatenated into single long memory, and one neuron encodes them. We consider this as hierarchization of episodic memories.

3.1 Encoding Episodic Memory

After a selection from stored episodic memories (this process will be explained at the section 3.3), the model starts to perform actions encoded in the episodic memory. When the model encounters an unexpected state, the model stops actions that are directed by the selected episodic memory, which is stored temporarily in the working memory module (Fig. 1b). At that time, the model makes a new episodic memory in the working memory module and decides whether to store it in the hierarchical episodic memory module by matching it with stored episodic memories. As an episodic memory, the model encodes a time series of states, actions, and reward or failure existence. Here, reward is given to the model at each goal location. Failure means that the model expected a reward but there is no reward.

3.2 Learning and Hierarchization of Episodes

Miyazaki and Kobayashi [16] claimed that no weight is effective to make a rational policy. We have followed their claim and excitatory connection between episodic memories which takes zero-one value. However, we use ten step values from 0 to 10 for inhibitory connection. When the model encounters an unexpected state or completes actions until the end of the selected episodic memory, learning and hierarchization of episodic memories happen. Here, we define "performed-E" as an episodic memory. A performed-E encodes the model's track which includes time series of positions and actions, and rewards. It corresponds to a selected episodic memory in case the model completes the selected episodic memory, or a new episodic memory in case the model encounter an unexpected state. In addition, "previous-E" indicates the previous performed-E, and "selected-E" indicates the selected episodic memory.

In learning, the model makes excitatory connection to the performed-E from the previous-E. If there is no reward in spite of the prediction according to the selected-E, the value of the inhibitory connection to the performed-E from the previous-E becomes 10. In other case, if the value of the inhibitory connection has more than 1 the model subtracts 1 from it.

If the performed-E or the previous-E has hierarchy, learning is taken into account in all excitatory connections of all the layers between both episodic memories. However, in case of inhibitory connection, it is taken into account in all connections from all the layers of the previous-E to only the top layer of the performed-E.

Moreover, the model makes excitatory connections from all the layers of the previous-E to the top layer of the selected-E. If there is no reward in spite of the prediction according to the selected-E, the value of the inhibitory connection from all the layers of the previous-E to the top layer of the selected-E becomes 10. In other case, if the value of the inhibitory connection between them has more than 1, the model subtracts 1 from it.

The model stores selected-Es up to 20 episodic memories in the working memory module until the model encounters an unexpected state. When the number of stored selected-Es exceeds 20, the working memory module refreshes it. In this research, hierarchization of episodic memories is realized with integrating continuous past selected-Es into one episodic memory, which is stored in the hierarchical episodic memory module. They are concatenated into single long memory, and one neuron encodes them. Thus, the hierarchization is performed when a failure or an unexpected reward occurred.

Excitatory and inhibitory connections from all the layers of the new episodic memory made by the hierarchization, to the top layer of the performed-E are constructed with the same rule in the section 3.3. At the hierarchization, the new hierarchized episodic memory becomes a performed-E if it ends with an unexpected reward. Otherwise, if the hierarchized episodic memory ends with failure, the hierarchized episodic memory is reconstructed in only successful episodic memories by dividing the last failure episodic memory. Thus, the last failure episodic memory becomes the performed-E.

3.3 Retrieval

The model chooses one of episodic memories connecting from the performed-E, which is stored in the hierarchical episodic memory module and shows the value of the inhibitory connection 0, based on number of rewards that each episodic memory has. A little of randomness is considered at this retrieval process.

Notice that the smallest episodic memory is state-action-state triplet. If there is no excitatory connection from the performed-E, or if there is no performed-E, the model chooses random action and then the model makes the smallest episodic memory as the next performed-E.

When the model cannot reach the goal within 300 decisions (retrievals), the model chooses random action.

4 Results

In this section, we compere our model implementing hierarchical episodic memory with non-hierarchical episodic memory. As non-hierarchal episodic memory, we restrict learning of the model. That is, connections from the previous-E to the performed-E and the selected-E, and from the hierarchized episodic memory to the performed-E, are limited only the top layer of each episodic memory. Fig. 2 and Fig. 3 show the results of the deterministic goal change described in the section 2.1 and 2.3.

Typical examples of the episodic memories' making process in the model are shown in Fig. 2. The x-axis is the number of decisions that the model chooses as one of the stored episodic memories. The y-axis is the index number of the performed-E, which means result of the selected-E. The models increased episodic memories, then suddenly stopped increasing and started to choose only a few episodic memories. It means that the models could predict all of task changes at correct timings.

Fig. 3 shows the learning curves. The x-axis is the number of Trial sets while the y-axis is the average number of failures on each 100 agents. In our previous research, the model without the hierarchization already realized prediction of long-term goal changes. In Fig. 3, we can confirm that the model with the hierarchization also learned to predict long-term goal changes even faster than the previous model. Moreover, after the partial pattern learning described in the section 2.3, the model learned faster than the model which learned only final pattern described in the section 2.1.

Figs. 4 to 6 show the results of the stochastic route change (see the section 2.2). Fig. 4 shows the typical examples of the episodic memories' making process in the model. The x-axis is the number of decisions that the model chooses as one of the stored episodic memories. The y-axis is the index number of the performed-E. The models with the hierarchization show slow or no increase in episodic memories (Fig. 4a), but all the agents without the hierarchization continued to keep increasing (refer to Fig. 4b as examples).

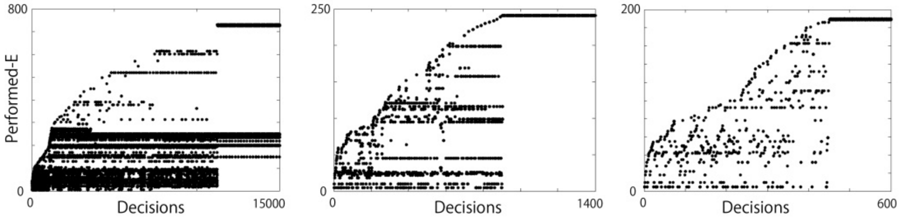

Fig. 2. Examples of the performed-E that the model made at the deterministic goal change task

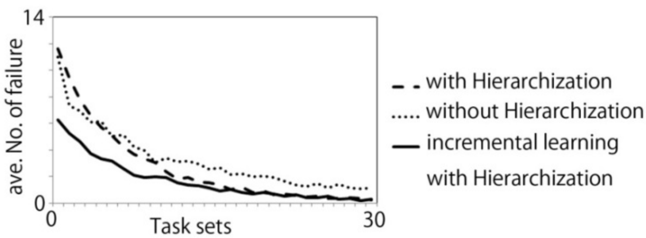

Fig. 3. Learning curves of Deterministic change (see the section 2.1 and 2.3)

Fig. 5 shows the learning curves. The x-axis is the number of Trial sets while the y-axis is the average number of failures on each 100 agents. The suggested model with the hierarchization showed better performance than our previous model. These slow or no increase in episodic memories (Fig. 4a) and better performance (Fig. 5) comes from the model's selection of episodic memories limited to some episodes.

Fig. 6 shows typical examples. The model with the hierarchization came to be limited to some episodic memories. Especially, in Fig.6 a2), we can see that the model divided the environment into 4 episodic memories. Then the model could limit the number of selected-Es (Fig. 6 a3).

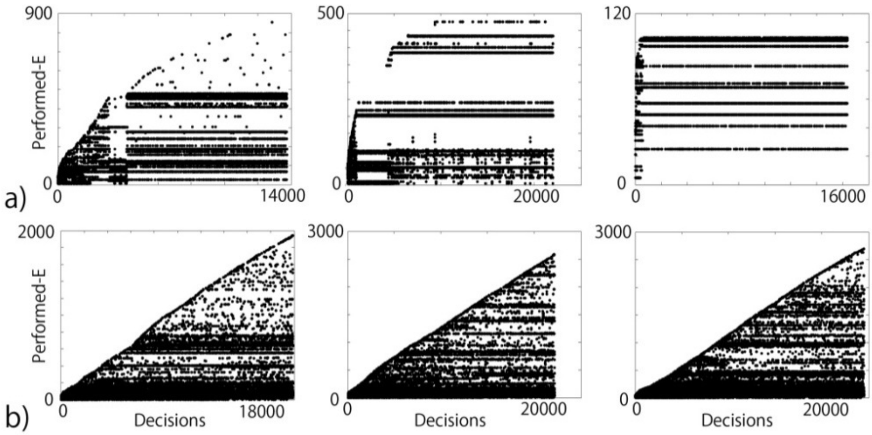

Fig. 4. Examples of Stochastic Route Change. a) is with the hierarchization. b) is without it.

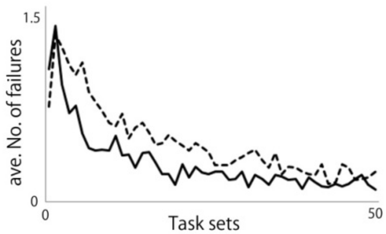

Fig. 5. Learning curves of Stochastic Route Change (see the section 2.2). The solid line shows with the hierarchization and the dashed line shows without the hierarchization.

5 Discussion

Recently, episodic future thinking [9] and episodic memory make attentions rapidly from many researchers. They consider episodes are necessary for future thinking. Although there is a deep relationship between both, its mechanism is still unknown.

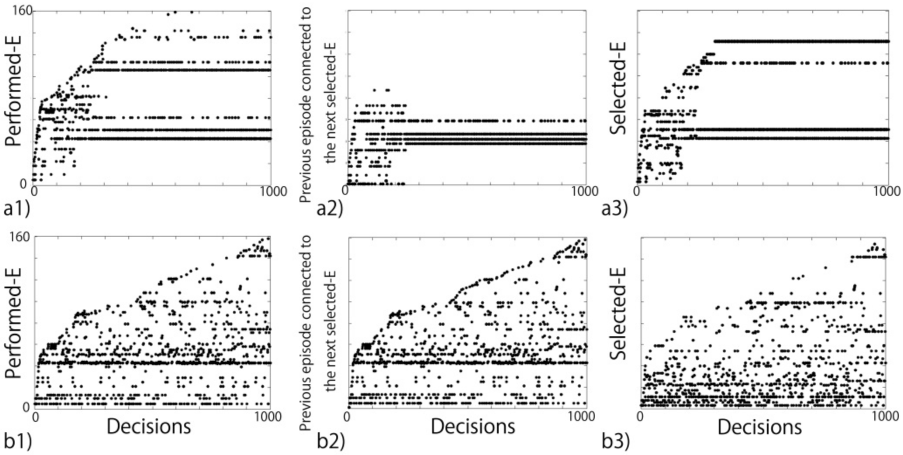

Fig. 6. Examples of the performed-E, Previous episode connected to the next selected-E, and the selected-E in the stochastic route change. a1-a3) are with the hierarchization, and b1-b3) are without the hierarchization. See also the section 2.2.

Nuxoll and Laird [14] tried to generalize episodic memory by implementing it in AI agent. They integrated their episodic memory module within the Soar cognitive architecture [17]. Because Soar is a general cognitive architecture that has been used to a variety of tasks, Nuxoll and Laird's model has a large possibility that various subjects will improve by episodic memory. However, as they claim, their model could not predict long-term environmental changes like task changes. We consider that simple addition of episodic memory is not enough.

In this paper, we have showed that our improved model with the hierarchization of episodic memories could predict long-term environmental change, realize incremental learning, and select adequate actions under POMDPs. It is expected that the same mechanism including editing and hierarchizing episodic memory is an underlying factor for the prediction of future events.

References

1. Tulving, E.: Elements of episodic memory. Oxford University Press, Oxford (1983)
2. Bond, A.H.: Representing episodic memory in a system-level model of the brain. Neurocomputing 65-66, 261–273 (2005)
3. Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., Tanila, H.: The hippocampus, memory, review and place cells: is it spatial memory or a memory space? Neuron. 23, 209–226 (1999)
4. Aota, Y., Miyake, Y., Ukai, S.: Neural network modeling of hippocampal place cells in rats. IEICE Transactions on Information and Systems J82-DII, 2355–2366 (1999)
5. Shapiro, M.L., Tanila, H., Eichenbaum, H.: Cues that hippocampal place cells encode: dynamic and hierarchical representation of local and distal stimuli. Hippocampus 7, 624–642 (1997)

6. Skaggs, W.E., McNaughton, B.L.: Spatial firing properties of hippocampal CA1 populations in an environment containing two visually identical regions. J. Neurosci. 18(20), 8455–8466 (1998)
7. Smith, D.M., Mizumori, S.J.Y.: Hippocampal place cells, context, and episodic memory. Hippocampus 16, 716–729 (2006)
8. Lipton, P.A., White, J.A., Eichenbaum, H.: Disambiguation of Overlapping Experiences by Neurons in the Medial Entorhinal Cortex. J. Neurosci. 27(21), 5787–5795 (2007)
9. Atance, C.M., O'Neill, D.K.: Episodic future thinking. Trends. Cogn. Sci. 5(12), 533–539 (2001)
10. Aota, Y., Yamaguchi, Y.: An algorithm for solving maze tasks with episodic memory integration. IPSJ Trans. Math. Model. Apps. 47(14), 93–107 (2006)
11. Aota, Y., Yamaguchi, Y.: Relationship between structure of tasks and past experiences on the learning algorithm with episodic memory integration. IPSJ Trans. Math. Model. Apps. 48(6), 12–22 (2007)
12. Zilli, E.A., Hasselmo, M.E.: Modeling the role of working memory and episodic memory in behavioral tasks. Hippocampus 18, 193–209 (2008)
13. Wickramasinghe, L.K., Alahakoon, L.D., Smith-Miles, K.: A novel episodic associative memory model for enhanced classification accuracy. Pattern Recognition Letters 28, 1193–1202 (2007)
14. Nuxoll, A., Laird, J.E.: Enhancing intelligent agents with episodic memory. Cogn. Sys. Res. 17-18, 34–48 (2012)
15. Brom, C., Lukavsky, J.: Towards virtual characters with a full episodic memory II: The episodic memory strikes back. In: Proc. Empathic Agents, AAMAS Workshop, pp. 1–9 (2009)
16. Miyazaki, K., Kobayashi, S.: Learning deterministic policies in partially observable markov decision processes. In: Proceedings of the 5th International Conference on Intelligent Autonomous System (IAS-5), pp. 250–257 (1998)
17. Laird, J.E.: Extending the Soar cognitive architecture. In: Proceedings of the First Artificial General Intelligence Conference, pp. 224–235 (2008)