Sequence Learning and Planning on

Associative Spiking Neural Network

Masayasu Atsumi
Department of Information Systems Science
Faculty of Engineering, Soka University
1-236 Tangi-cho, Hachioji-shi, Tokyo 192-8577, Japan
matsumi@t.soka.ac.jp

Abstract - We have been building an auto/hetero-associative spiking neural network combined with a working memory model. In this model, a state-driven forward sequence and a goal-driven backward sequence on the associative network are respectively represented by a sequence of synchronous firing in a particular gamma subcycle during a theta oscillation. These forward and backward sequence firings are transmitted to the working memory, temporarily maintained, and integrated based on a competition principle to make a plan. This paper shows that our system can learn forward and backward sequences simultaneously and a plan is incrementally synthesized by repeating their recall and integration.

I. INTRODUCTION

Working memory is an active system necessary for execution of goal-seeking cognitive tasks and it provides temporary storage and manipulation of information retrieved from a long-term memory [1], [2]. Planning is considered as one of tasks for which working memory is necessary. We assume that planning is executed in working memory through integrating a state-driven forward sequence association with a goal-driven backward sequence association under attention control. As one of neural network models of sequence association, Jensen and Lisman proposed a biologically plausible auto/hetero-associative network model [3], [4], in which an auto-associative network in the cortex encodes individual items and hetero-associative network in the hippocampus encodes their forward sequence that is time series of their occurrences. However in our assumption, a goal-driven backward sequence, that is retrospective time series of occurrences, must be involved in planning process in addition to a state-driven forward sequence.

We have been building an auto/hetero-associative spiking neural network combined with a working memory model [5], in which a state-driven forward sequence and a goal-driven backward sequence on the associative network are integrated in the working memory to make a plan. In this paper, we extend the sequence integration

algorithm proposed in [5] and then show that our system can learn forward and backward sequences simultaneously and a plan is incrementally synthesized through repetition of their recall and integration, by using discrete pulse-driven neural network simulations of a route planning problem.

II. THE MODEL

A. Representation

In our model, a state-driven forward sequence and a goal-driven backward sequence are respectively represented by a sequence of synchronous firing in a particular gamma subcycle during a theta oscillation (Fig. 1).
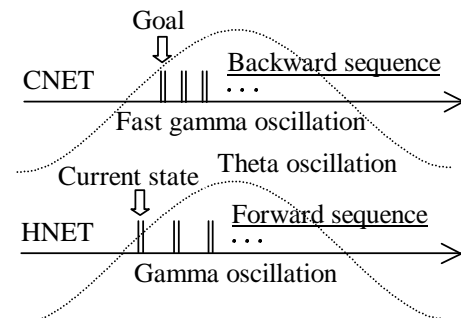

Fig. 1. Sequence representation.

This representation is supported by findings that the hippocampus and some regions of cortex involved in memory storage show dual oscillations in which a low-frequency theta oscillation is subdivided into subcycles by a high-frequency gamma oscillations, and that each different information is encoded by a subset of neurons that fire synchronously and stored in different gamma subcycles [3], [4]. For example, it is suggested that dual oscillations serve as a multiplexing mechanism by which multiple short-term memories can be actively maintained [6]. Also it is discovered that a firing sequence of place cells during gamma-frequency recall corresponds to a prediction of expected positions during rat's move-

ment over time [7].

In planning in our model, these forward and backward sequence firings are transmitted to working memory, temporarily maintained, and integrated based on a competition principle.

B. Neural Network Architecture

The neural network architecture is shown in Fig. 2. The network is composed of the auto/hetero-associative network "CNET" that is a model of the cortex, the hetero-associative network "HNET" that is a model of the hippocampus, and the working memory.
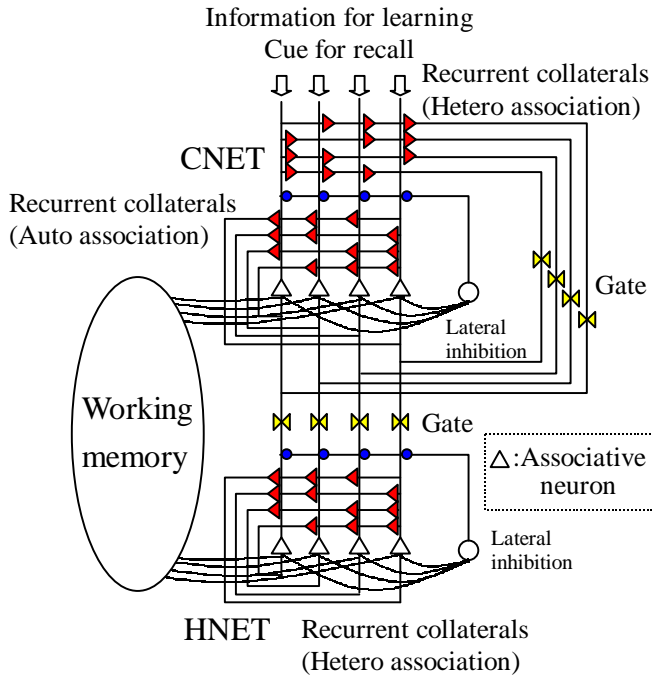

Fig. 2. The neural network architecture: SEQAINET (SE-Qence Association and Integration NETwork). Parameters applied to the CNET are as follows. As for auto-associative recurrent collaterals, $\tau_{ax} = 2$ msec, $\tau_{sr}^{+} = \tau_{sr}^{-} = 2$ msec, and $\tau_{sp} = 1$ msec. As for hetero-associative recurrent collaterals, $\tau_{ax} = 10$ msec, $\tau_{sr}^{+} = 2$ msec, $\tau_{sr}^{-} = 20$ msec, and $\tau_{sp} = 10$ msec. As for lateral inhibition, $w_I = 0.9$ for learning and 0.4 for recall, $\tau_{sp} = 14$ msec for learning and 2 msec for recall, and $\tau_{ax} = 2$ msec. Parameters applied to the HNET are as follows. As for hetero-associative recurrent collaterals, $\tau_{ax} = 20$ msec, $\tau_{sr}^{+} = 20$ msec, $\tau_{sr}^{-} = 2$ msec, and $\tau_{sp} = 30$ msec. As for lateral inhibition, $w_I = 0.9$ for learning and 0.8 for recall, $\tau_{sp} = 12$ msec for learning and 10 msec for recall, and $\tau_{ax} = 4$ msec. For all associative neurons in the CNET and the HNET, $\tau_{abs} = 125$ msec and $\theta = 0.45$. A conduction time from the CNET to the HNET $\tau_{ax} = 1$ msec and $\lambda = 0.9$.

The principal excitatory spiking neuron in associative networks is formulated as follows. We call this neuron an associative neuron. Let t_i^f be a recent firing time of a neuron i before time t and t_{ij}^a be a recent time of pulse arrival from a neuron j to i before time t. The membrane potential $p_i(t)$ of a neuron i at time t is given by

$$p_i(t) = \frac{ep_i(t)}{N_i(t)} + ext_i(t) - ip_i(t) - r(t - t_i^f) \qquad (1)$$

where $ep_i(\cdot)$ is the sum of EPSPs of recurrent collaterals, $ip_i(\cdot)$ is the sum of IPSPs of lateral inhibition, $ext_i(\cdot)$ is an external input, $r(\cdot)$ is the refractory function, and $N_i(\cdot)$ is a normalization factor. The $ep_i(t)$ is given by

$$ep_i(t) = \sum_j ep_{ij}(t) = \sum_j (kp_{ij}(t) \times w_{ij}(t)) \qquad (2)$$

where $ep_{ij}(\cdot)$ is an EPSP and $w_{ij}(\cdot)$ is synaptic efficacy from a neuron j to i. The $kp_{ij}(t)$ is a kernel function given by

$$kp_{ij}(t) = kp_{ij}(t_{ij}^a) \times \exp(-\frac{t - t_{ij}^a}{\tau_{sp}}) + \delta_{ij}(t) \qquad (3)$$

where τ_{sp} is a time constant of synaptic potential and the $\delta_{ij}(t)$ is a function which takes a value of 1 when a pulse arrives from a neuron j to i at time t, otherwise 0. The $ip_i(t)$ is given by

$$ip_i(t) = \sum_j ip_{ij}(t) = \sum_j (kp_{ij}(t) \times w_I) \qquad (4)$$

where $ip_{ij}(\cdot)$ is an IPSP from a neuron j to i and w_I is a constant that represents non-plastic synaptic efficacy. The refractory function $r(\cdot)$ takes positive infinity for $t - t_i^f \leq \tau_{abs}$ and otherwise 0, where τ_{abs} is the absolute refractory period. The $N_i(t)$ is calculated by $N_i(t) = \sqrt{\sum_j kp_{ij}(t)^2}$. Since transmission of a pulse takes a conduction time τ_{ax}, the relation $t_{ij}^a = t_j^f + \tau_{ax}$ holds. An associative neuron fires when $p_i(t) \geq \theta$ is satisfied, where θ is a threshold, and then afterhyperpolarization is applied.

Synaptic efficacy of excitatory recurrent collaterals is modulated by the following Hebbian rule.
(1)when a pulse arrives from a presynaptic neuron j to a postsynaptic neuron i at time t, the synaptic efficacy $w_{ij}(t)$ is modulated according to

$$\Delta w_{ij}(t) = \lambda \times W(t_i^f - t) \times (1 - w_{ij}(t)), \qquad (5)$$

(2)and if a postsynaptic neuron i fires at time t when a pulse arrives from presynaptic neuron j, all synaptic

efficacy w_{ik} on all recurrent collaterals are modulated according to

$$\Delta w_{ik}(t) = \lambda \times W(t - t_{ik}^a) \times (\frac{kp_{ik}(t)}{N_i(t)} - w_{ik}(t)). \quad (6)$$

In above formulas, λ is a modulation rate and the $W(s)$ is a modulation window given by

$$W(s) = \begin{cases} exp(-\dfrac{s}{\tau_{sr}^+}) & (s \geq 0) \\ exp(\dfrac{s}{\tau_{sr}^-}) & (s < 0) \end{cases} \quad (7)$$

where τ_{sr}^+ and τ_{sr}^- are time constants of synaptic modulation.

Main parameters that characterize associative networks are a time constant of synaptic modulation for recurrent collaterals, synaptic efficacy and a time constant of synaptic potential for lateral inhibition, a conduction time, and the absolute refractory period of associative neurons. Dependent on different parameter setting, the CNET acquires feature of auto-association and backward hetero-association and the HNET acquires feature of forward hetero-association, as described in section III. In our setup, forward hetero-association in the HNET occurs on gamma cycles and represents a state-driven forward sequence, and backward hetero-association in the CNET occurs on shorter cycles than firing in the HNET and represents a goal-driven backward sequence. We call this cycle a short gamma cycle.

C. Sequence Maintenance and Integration

Principal functions for planning in working memory are sequence maintenance and sequence integration.

In sequence maintenance, a firing sequence $C_0, C_1, \ldots, C_n(\equiv \{C_i\})$ of a gamma cycle whose first pulse pattern C_0 encodes a given current state and a firing sequence $G_0, G_1, \ldots, G_m(\equiv \{G_j\})$ of a short gamma cycle whose first pulse pattern G_0 encodes a given goal are maintained by filtering firing sequences transmitted to working memory from the HNET and the CNET. The $\{C_i\}$ represents a sequence of pulse patterns driven by the current state and the $\{G_j\}$ represents a sequence of pulse patterns driven by the goal. For convenience' sake, C_i and G_j are assumed to be Boolean arrays, an element of which is $true$ for firing, otherwise $false$. At each time a new goal-driven pulse pattern G_k is filtered, it is checked whether $C_1 \wedge G_k \neq F$, where F is a Boolean array all elements of which are $false$. This condition means that the first pulse pattern C_1 in forward firing and the pulse pattern G_k in backward firing contain representation of the same object, which

suggests that the backward firing may be suppressed at that time because backward firing after that time is irrelevant to integrating a plan. When this condition holds, the CNET is temporarily inhibited by feedback not to propagate irrelevant firing, which is exemplified in section IV.

Forward and backward sequences $\{C_i\}$ and $\{G_j\}$ are maintained for a theta period and sequence integration is executed on them at the end of the theta period. In sequence integration, for each pulse pattern of the state-driven firing sequence $C_1, C_2, \ldots, C_n$, the following operation is executed between C_i and the goal-driven firing sequence $G_0, G_1, \ldots, G_m$ in the order of firing time $i = 1, 2, \ldots, n$. Let T be the maximum index of pulse patterns to be operated in $\{G_j\}$. At the start, $T := m$.

[Goal-directed competitive selection] For each pulse pattern of a goal-driven firing sequence $G_0, G_1, \ldots, G_T$, $P_{ij} := C_i \wedge G_j$ is computed in the order of firing time $j = 0, 1, \ldots, T$ and the first $P_i(\equiv P_{iJ})$ all elements of which are not $false$ is obtained. If such a P_i is found, it is regarded as a winning pulse pattern that constitutes a plan. □

If P_i equals to G_0, sequence integration terminates. Otherwise, let $T := J$ and continue the operation.

This operation is logically interpreted as selection of a sequence of pulse patterns in forward firing that intersect backward firing. A firing sequence obtained by applying the goal-directed competitive selection to $C_1, C_2, \ldots, C_n$ is regarded as a plan.

D. Problem Specification and an Example

In this paper, we focus on problems in which sequences are correlated but each pulse pattern, which is an element of a sequence, is mutually orthogonal. As an example, we use a route planning problem in which a robot is about to make a plan to reach a goal.
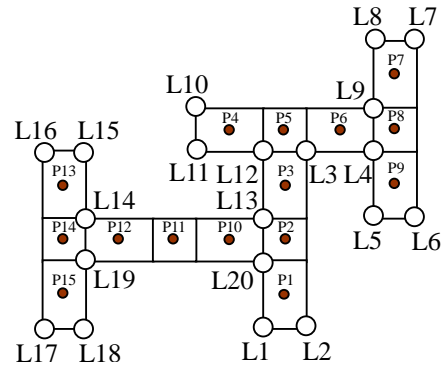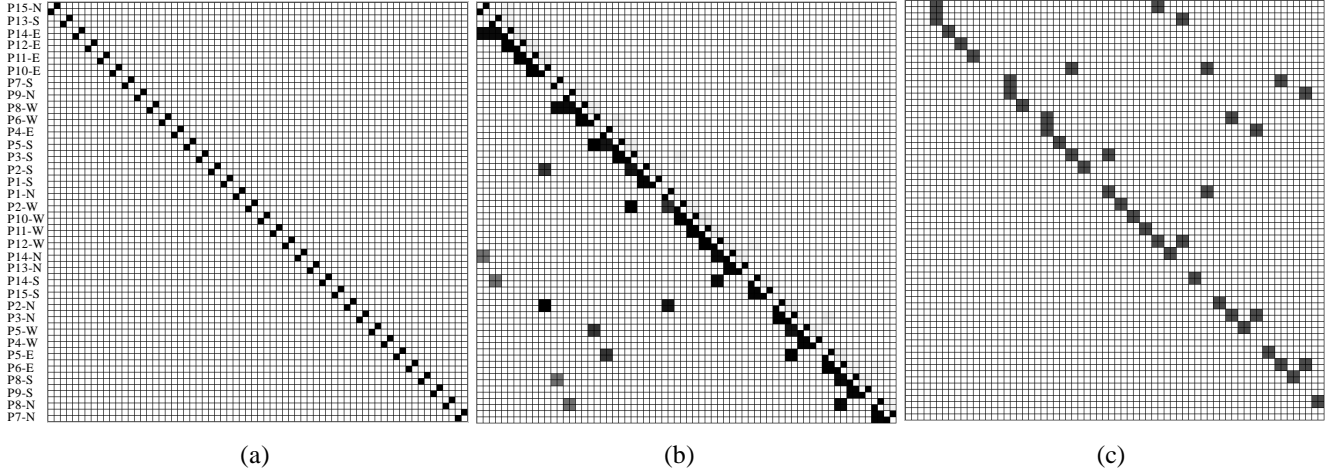

Fig. 3. T-maze as an example.

Fig. 4. Synaptic efficacy matrices of (a)inner recurrent collaterals of the CNET (auto-association), (b)outer recurrent collaterals of the CNET (backward hetero-association), and (c)recurrent collaterals of the HNET (forward hetero-association). A "Pi-j" denotes a pair of associative neurons which corresponds to a view at a point "Pi" toward the "j"-direction ∈ {North, South, East, West}.

Fig. 3 shows a T-maze used in experiments. The T-maze has 20 landmarks L1-L20. The robot pays attention to a nearby landmark in right and left walls respectively at each point P1-P15. Since a landmark can be seen in one of four directions in each side, we can encode a view of the robot by using 160 bits that is the number of landmarks multiplied by the number of directions. As a result, a view of the robot at each point is expressed by a code of 160 bits whose two bits are on and others are off, which satisfies each pulse pattern for a view is mutually orthogonal but sequences of views are correlated. This code corresponds to an input pulse pattern to the network, and the CNET and the HNET have 160 associative neurons respectively since they memorize this view information.

III. SEQUENCE LEARNING

Each input for learning is encoded to a pulse pattern and repeatedly inserted in theta cycles. Pulse patterns for different inputs are inserted in different gamma subcycles of the theta cycle. In experiments, a theta cycle is 250 $msec$ and a gamma cycle is 20 $msec$. Each pulse pattern that encodes a view at each point consists of synchronous firing of two neurons and is inserted five times in a way of partly overlapping with firing of a preceding pulse pattern. For example, a pulse pattern for a view at P1 is inserted at 0, 250, 500, 750, and 1000 $msec$, and a pulse pattern for a view at P2 next to P1 is inserted at 770, 1020, 1270, 1520, and 1770 $msec$.

Fig. 4 shows synaptic efficacy matrices of recurrent collaterals of the CNET and the HNET after 30 routes from each end point P1, P9, P7, P4, P13, and P15 to all other end points of T-maze have been learned once in random order. In these matrices, reinforcement of upper right elements of the diagonal means that forward hetero-association is acquired, and reinforcement of lower left elements of the diagonal means that backward hetero-association is acquired. We can see the CNET acquires auto-association of input patterns and backward hetero-association of them, and the HNET acquires forward hetero-association of them. Synaptic efficacy for forward hetero-association and backward hetero-association exhibits symmetry.

Requirements for the CNET to acquire these features are as follows. As for auto-association, it is necessary that a conduction time on recurrent collaterals is short enough in comparison with a gamma cycle, and a modulation window of synapses on recurrent collaterals should be short that synaptic efficacy is reinforced only when a pulse arrival and postsynaptic firing rigidly synchronize in a short interval. As for backward hetero-association, it is necessary that a conduction time on recurrent collaterals is longer than the one for auto-association but shorter than the gamma cycle, and a modulation window of synapses on recurrent collaterals should satisfy a condition that synaptic efficacy is reinforced even when a pulse arrives with a long delay after postsynaptic firing. Additionally, in order not to confuse auto- and hetero-association, a strong and lasting lateral inhibition and a long period of absolute refractoriness are needed. Requirements for the HNET are as follows. A conduction time on recurrent collaterals should be nearly equal to a

gamma cycle and a modulation window of synapses on recurrent collaterals should satisfy a condition that synaptic efficacy is reinforced even when firing occurs with a long delay after a pulse arrival. Additionally, in order not to confuse different hetero-association, a strong and lasting lateral inhibition and a long period of absolute refractoriness are needed. Parameters used in experiments to satisfy these requirements are shown in Fig. 2.

IV. GOAL-DIRECTED RECALL FOR PLANNING

In planning, a current state and a goal are given as a goal-directed recall cue. That is, pulse patterns that encode a current state and a goal are inserted into the network in a certain interval.

To control goal-directed recall, gate control and lateral inhibition adjustment are executed as follows. As for gate control, gates between the CNET and the HNET concerned in transmitting firing of the CNET for a goal cue are closed, and gates on hetero-associative recurrent collaterals of the CNET concerned in transmitting firing for a current state cue are closed (Fig. 2). As a result, pulses that originate in the current state cue are propagated only to the HNET and pulses that originate in the goal cue are propagated in the CNET. As for lateral inhibition adjustment, lateral inhibition of the CNET and the HNET is weakened to activate successive hetero-associative firing, which was suppressed at learning by strong lateral inhibition to competitively emphasize a sequence to be learned. Concretely in experiments, synaptic efficacy w_I and a time constant of synaptic potential τ_{sp} for lateral inhibition are modified as shown in Fig. 2. The balance of lateral inhibition between the CNET and the HNET is closely related to our way of integration of forward and backward sequence recall. We adjust lateral inhibition so that backward sequence is broadly propagated but propagation of forward sequence is restricted to be narrow. The effect of this adjustment is explained in experimental results below.

Planning is executed through repeating forward and backward sequence recall in the HNET and the CNET and their maintenance and integration in working memory during subsequent theta cycles, in which a plan is incrementally recognized. Fig. 5 shows backward sequence firing in the CNET, forward sequence firing in the HNET, and a result of their integration for planning of a route from a point P13 to a goal P9, which is drawn with bold line, after all routes in the T-maze have previously been learned. A pulse pattern for a view at P13 is inserted at 0 $msec$ as a current state cue and a pulse pattern for a view at P9 is inserted at 14 $msec$ as a goal cue. Each pulse pattern that encodes a view at each point consists
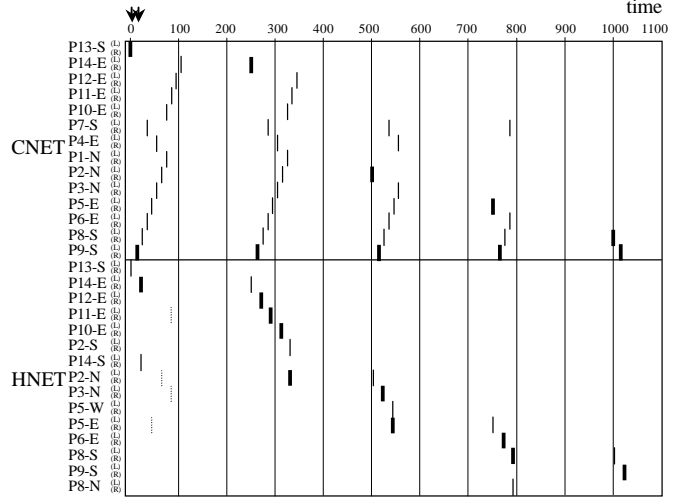

Fig. 5. Goal-directed recall and integration for planning. A denotation of "Pi-j" is the same as that in Fig. 4. "L" denotes the left view and "R" denotes the right view. Bold bars in the CNET part express firing of a state cue and a goal cue. Bold bars in the HNET part express winning firing by the goal-directed competitive selection in working memory. This firing sequence of bold line represents a plan.

of synchronous firing of two neurons. We can observe pulse patterns on gamma cycles of 20 $msec$ in the HNET and pulse patterns on short gamma cycles of 10 $msec$ in the CNET.

In the first theta cycle, pulse patterns in the CNET from 14 $msec$ to 104 $msec$ represent a backward sequence recall of views toward the goal P9 on all learned routes. That is, a view at P9 is recalled at 14 $msec$, a view at P8 toward P9 is recalled at 24 $msec$, two views at P6 and P7 toward P8 are recalled at 34 $msec$, and in the same way backward recall of views toward the goal occurs in succession. In this sequence, we can see that firing after 104 $msec$, which corresponds to a view at P15 toward P14, is suppressed and does not occur. This is caused by feedback inhibition imposed by sequence maintenance in working memory as described in section II. In the HNET, a pulse pattern that encodes a view at the current point P13 is observed at 1 $msec$ and two pulse patterns that encode views at a fork P14 next to P13 toward P12 and P15 are observed at 21 $msec$, but firing does not continue thereafter in this gamma cycles[1]. This represents that only a part of forward sequences of views on learned routes is recalled. The cause of firing extinction at a fork is explained as follows. When

---

[1]Pulse patterns drawn with dotted line occur in another gamma cycles with phase shift.

plural views are recalled, the number of associative neurons that fire increases, which makes lateral inhibition stronger than when one view is recalled, causing firing extinction. As a result of sequence integration in working memory, it is observed that one pulse pattern for a view at P14 toward P12 is selected. This represents a sub-route P13-P14 and a direction to P12 are planned to reach P9. This result means that when forward sequence recall can have several branches, their recall does not continue in parallel but one of them is selected by attention control based on the goal-directed competition in sequence integration. By the way, we can see that forward sequence firing stops at a fork but backward sequence firing does not. This is because of our lateral inhibition adjustment as described above, in which lateral inhibition in the HNET is stronger than that in the CNET, that is, competition in forward sequence recall is stronger than that in backward sequence recall. Non-parallel forward sequence recall and the goal-directed attention control at a branch point depend on this adjustment.

Planning continues in the second theta cycle by introducing the winning pulse pattern into the network as if it is a state cue. We can see in the HNET that a forward sequence of views is recalled from 251 $msec$ to 311 $msec$, two views are recalled at 331 $msec$, and one of these views is selected through sequence integration. As a result, a route P14-P12-P11-P10-P2 and a direction to P3 are recognized as a sub-plan. In subsequent theta cycles, planning continues in the same way and the whole route is incrementally planned, that is, the plan is extended to a sub-route P2-P3-P5 in the third theta cycle, a sub-route P5-P6-P8 in the forth theta cycle, and a sub-route P8-P9 in the fifth theta cycle.

V. CONCLUSION

We have presented the auto/hetero-associative spiking neural network combined with a working memory model, in which a state-driven forward sequence and a goal-driven backward sequence on the associative network are integrated in the working memory to make a plan. By simulation experiments, we have confirmed firstly that our associative network can learn forward sequence and backward sequence simultaneously. Secondly, it has been confirmed that a plan is incrementally synthesized by repeating forward and backward sequence recall on associative networks and their integration in working memory during subsequent theta cycles. Especially, it was found that the goal-directed competition performed attention control for selecting one of several branches in planning.

References

[1] A. Baddeley, "Working Memory," *Oxford University Press*, 1986.

[2] A. Baddeley, "Working Memory," *Science*, 255, pp.556-559, 1992.

[3] O. Jensen and J. E. Lisman, "Physiologically Realistic Formation of Autoassociative Memory in Networks with Theta/gamma Oscillations: Role of Fast NMDA Channels," *Learning & Memory*, 3, pp.243-256, 1996.

[4] O. Jensen and J. E. Lisman, "Theta/gamma Networks with Slow NMDA Channels Learn Sequences and Encode Episodic Memory: Role of NMDA Channels in Recall," *Learning & Memory*, 3, pp.264-278, 1996.

[5] M. Atsumi, "Sequence Memories and their Integration for Planning: A Spiking Neural Network Model," *ICONIP 2001 Proceedings*, pp.891-896, 2001.

[6] J. E. Lisman and M. P. A. Idiart, "Storage of 7±2 Short-term Memories in Oscillatory Subcycles," *Science*, 267, pp.1512-1515, 1995.

[7] J. O'Keefe and M. L. Recce, "Phase Relationship between Hippocampal Place Units and the EEG Theta Rhythm," *Hippocampus*, 3(3), pp.317-330, 1993.