Approach to Integrate Episodic Memory into Cogency-based Behavior Planner for Robots

Min-Joo Kim, Seung-Hwan Baek, Se-Hyoung Cho, and Jong-Hwan Kim
Robotics Program, School of Electrical Engineering, KAIST
291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea
{mjkim, shbaek, shcho, johkim}@rit.kaist.ac.kr

***Abstract*—This paper proposes a novel scheme of integrating episodic memory into semantic memory based task planner. Task planners have taken an important role in AI research along with semantic memory to better perform tasks for robots. Episodic memory memorizes and retrieves temporal sequence of situated behaviors by which temporal relationship between behaviors can be defined. None of any research, however, has implemented it into their work for task planning. By introducing episodic memory into task planner, the temporal causal relationship between situated behaviors, which are stored in semantic memory, is taken into consideration. The integrated architecture proves its effectiveness by notably reducing the number of nodes traversed in finding solutions. Robots can reduce time complexity in solving given problems by retrieving previous memories. Deep Adaptive Resonance Theory (Deep-ART) neural model and cogency-based hierarchical behavior planner are used for the episodic memory and the task planner, respectively. Cogency-based hierarchical behavior planner proves its capability of solving given problems in experiment with humanoid robot Mybot, and Deep-ART is augmented to the planner and tested in simulations. Therefore, the contribution of this approach lies on developing a framework which takes advantage of implementing episodic memory and planner in one place.**

I. Introduction

Various types of task planners have been developed to perform the tasks given to robots. In their applications to robot for intelligent motions, memory takes an important role. The most frequently used memory is semantic memory which provides objects, its usage and possible behaviors associated with them. With the information provided by the semantic memory, planners search for possible sequence of behavior solutions. However, while searching for the most likely behaviors, it has to traverse less likely behaviors. The episodic memory helps in the context of suggesting the most relevant events, to reduce time complexity of behavior planning from scratch; higher level planning has more chance to match with plans in episodic memory. Most importantly, the episodic memory can provide temporal relationship between situated behaviors and the behavior planner.

There have been significant research done on the episodic memory. An adaptive resonance architecture (ART 1) introduced an unsupervised learning models for binary inputs [1], while fuzzy adaptive resonance theory (fuzzy ART) allowed continuous inputs [2]. Based on these works, a memory architecture model to memorize and retrieve spatio-temporal information of episodes has been proposed in terms of episodic memory adaptive resonance theory (EM-ART) [3]. Deep-ART has been proposed to enhance EM-ART's performance. Also, a novel approach to describe and mimic neocortex's structural and algorithmic attributes was developed by the name, hierarchical temporal memory (HTM) [4]. This HTM model allows to process enormous amount of input patterns to be learned and memorized in temporal fashion so that it enables anomaly detection in current situation compared to the previous ones.

Task planner is a hierarchical architecture which changes the initial state of environment to the desired target state by a temporal sequence of required robot behaviors on proper objects. In a recent study, common sense domain knowledge as a semantic knowledge was integrated with task planner to improve semantic relationship between objects and places in environment [5]. Memory-based architecture is necessary to realize task intelligence [6]. Also, while normal task planners use shallow domain knowledge, an approach proposed to integrate hierarchical spatial information with semantic knowledge allowing a robot advanced autonomy and intelligence [7]. Behavior planning can be described by STRIPS which uses tree search algorithm [8]. Importance of learning in emergence of behaviors is described in [9]. Confabulation as heuristic measure in planner can be used in tree searching algorithm to mimic human intelligence [10].

This paper is organized as follows. Section II presents brief explanation about episodic memory structure. Section III describes development of cogency-based hierarchical behavior planner and Section IV explains the structure and logic of the episodic memory augmented task planner. Section V describes the simulation in MATLAB, results and discussion for Section VI. Finally, concluding remarks follow in Section VII.

II. Implemented Episodic memory architecture

Deep-ART is applied to implement an episodic memory with semantic memory based task planner. In the proposed architecture, five layered Deep-ART is designed to memorize the hierarchical sequences of behaviors (Fig. 1). The first channel of input field gets behaviors as inputs and the rest channels take objects, which is related to the behaviors, as inputs. Each input channel is linked to attribute field to store the relationship between types, instances, situation types, and situation instances of objects. The activated nodes in input field goes through matching process with the nodes in primitive behavior field. The input values which successively come into input field activate the nodes in primitive fields in temporal order to generate primitive behavior sequences. The primitive behavior sequence is entered to the manipulation behavior field to activate nodes of the field and generates sequences of manipulation behaviors. This process is repeated up to the task field level and enables memorization of hierarchical behaviors. The memorized hierarchical behaviors are retrieved to perform given tasks.

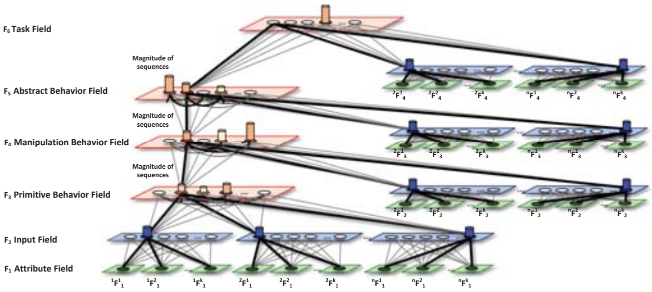


Figure 1: Deep-ART is used in the proposed memory augmented planner architecture.

III. Layered Confabulation based Behavior Selection

This section represents the behavior selection part in our proposed episodic memory augmented task planner. The planner used is a STRIPS based behavior selection method which draws goal-directed behavior sequences through tree search and an algorithm which merges cogent confabulation replicating human intelligence which automatizes frequently occurring behaviors. Then, it proposes a method which hierarchically plans the behavior sequences achieving objectives by applying this algorithm to a hierarchically defined behavior set.

The layered confabulation architectures has been developed for hierarchical behavior planning and proved its effectiveness on two scenarios: beverage serving, cereal and milk serving scenarios [11] and is shown in Fig. 2. Planning module receives from perception module the environmental information by situated affordance [12]. Task recognition module recognizes initial and goal states of task problem to solve and returns them to the planning module.

Planning module searches for a solution that satisfies the initial and goal states in a hierarchical manner. In the specific scenarios of beverage serving, cereal and milk serving with Mybot robot developed in the RIT lab at KAIST, the hierarchy of layered confabulation architecture consists of three layers: abstract level, manipulation level, and primitive level. Each layer plans a sequence of behaviors that satisfies pre-condition and post-condition returned from one level higher than the layer itself, and the process module which calculates the score to select the most proper behavior for pre-condition is the confabulation process and memory module.

In confabulation module, quantification from STRIPS is applied to express the relationship between situated objects and behaviors instances in probability. Situated objects and behaviors are the assumed fact symbols and conclusion symbols for semantic memory, respectively (Fig. 3). Situational context or status is understood by the situated objects. The most suitable behaviors in the situational context are decided and ranked by considering the cogency value by Bayes theorem (Eq. (5)). The following cost function is the modified version of traditional A* search algorithm and includes the empirically adjusted cogency ranking term:

$$f = g + h + (c_m - 1) \times k \times w, \tag{1}$$

with,

$$w = max(1 - \frac{n}{N}, 0), \tag{2}$$

$$c_m = \frac{Ranking(cogency(b_m))}{M}, \tag{3}$$

$$cogency(b_m) = \prod_{i=1}^{I} \prod_{j=1}^{J_i} p(< o_i, S_j^i > | b_m) \tag{4}$$

$$p(abcd|e) \approx k \cdot [p(a|e) \cdot p(b|e) \cdot p(c|e) \cdot p(d|e)] \tag{5}$$

where
g: one per each behavior
h: number of situated object symbols not matching with the goal state
k: gain constant of cogency ranking score
N: anticipated solution length of search tree
M: number of situated behaviors
n: depth of current node in search tree
$Ranking()$: cogency ranking
b_m: m^{th} behavior
$cogency(b_m)$: cogency value of m^{th} behavior
I: number of object instances over all of the object types
J_i: number of situation types of object instance o_i
S_j^i: substitution of s_k^{ij} with proper k, according to the current environmental state
$< o_i, S_j^i >$: situated object.

The f value returned to the planning module is used as metric to select proper behaviors for given pre-conditions

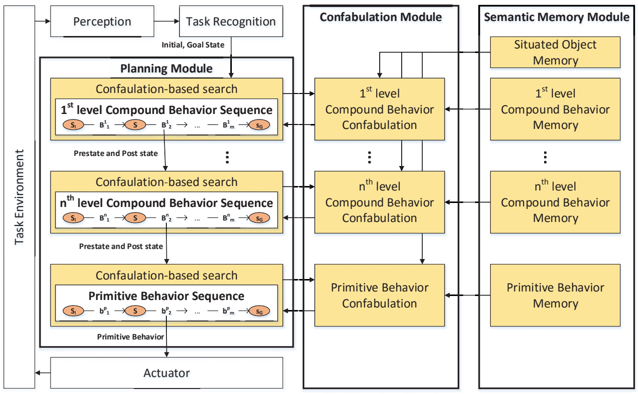


Figure 2: Overall behavior planning architecture

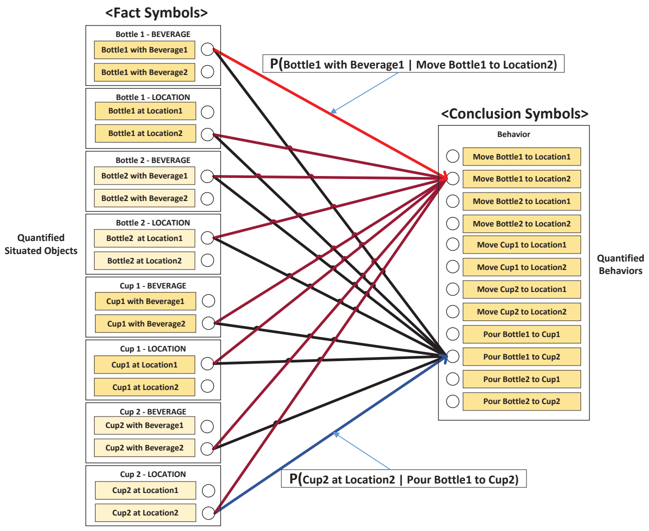


Figure 3: fact symbol and conclusion symbol in calculation of confabulation probability by Bayes theorem

(Fig. 4). As shown in the overall architecture figure, this process is performed in hierarchical order. The sequence of primitive behaviors, which are in the lowest level, are found and returned to the actuator for movement of robot. No0t that the effectiveness of this architecture was verified through experiments with Mybot (Fig. 5).

IV. Episodic Memory-Augmented Cogent Confabulation Behavior Planner

In this section, the process of integration of cogent confabulation behavior planner and episodic memory is described. The overall architecture is presented in Fig. 6.

A. Architecture

Compared to Fig. 2, ART-based Episodic memory part and Behavioral-causal Degree (BD) module in semantic memory module are additionally attached to the cogent confabulation behavior planner frame. The roles of ART-based Episodic

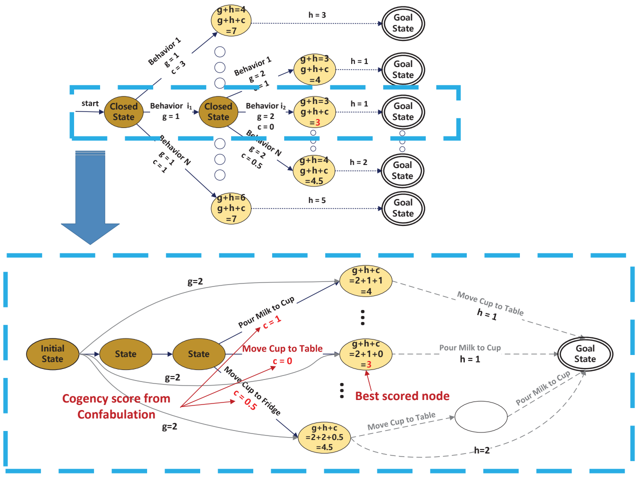


Figure 4: Behavior selection mechanism in behavior search algorithm.

(a) Initial state of experiment with Mybot

(b) Goal state of experiment with Mybot

Figure 5: Cogent confabulation planner experiment with Mybot for beverage serving task.

memory part are twofold; providing a default hierarchical sequence of behaviors that are retrieved by recognizing objects caught in vision of the robot, and providing a behavioral causal degree between frequently occurring behaviors. To match with the hierarchy of planner and semantic memory, ART-based episodic memory has three layers: abstract level, manipulation level, and primitive level. ART-based episodic memory retrieves the temporal sequence of behaviors in its memory for all the levels.

Behavioral-causal degree is employed to reinforce the causal relationship between a pair of two behaviors (Fig. 7). $^{o}F_m$ means the m^{th} level output field in ART architecture. d^m_{ki} means the behavioral-causal degree from the k^{th} behavior to i^{th} behavior of the m^{th} ART field. As the d^m_{ki} gets higher (the higher the better), the chance that the i^{th} behavior occurs after the k^{th} behavior does increases, i.e. the degree between two behaviors is short. Regarding structure of ART, refer to Fig. 1. The causal degree d is a directional value, therefore, for one pair of the k^{th} and i^{th} behaviors, there exist two causal degree values: d^m_{ki} and d^m_{ik}. A value for one direction does not necessarily affect the value for other direction. When a sequence of behaviors are retrieved by episodic memory, the BD values between all the pairs of

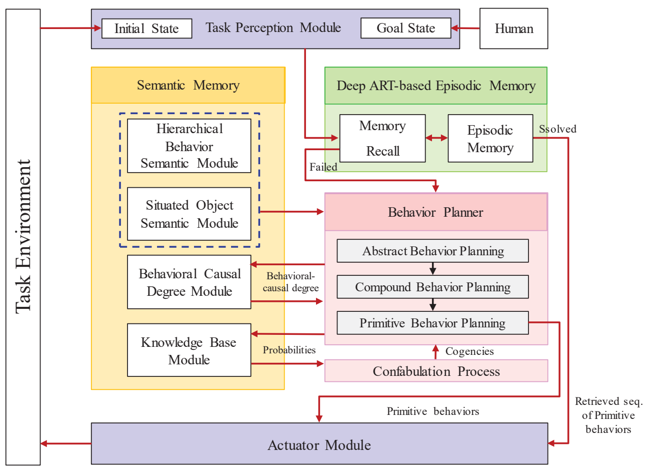


Figure 6: Overall architecture of Deep-ART augmented cogent confabulation behavior planner.

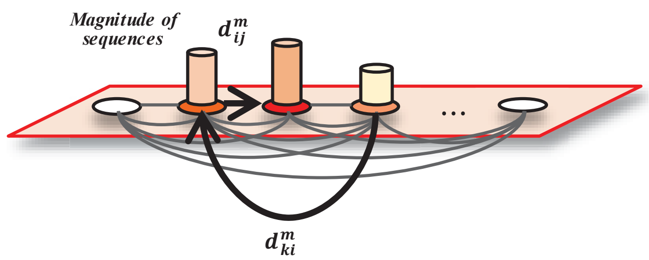


Figure 7: Behavioral-causal degree d^m of output channel in the m^{th} field of ART based episodic memory.

successive behaviors are returned together.

Because current environmental situations cannot perfectly match with previous ones in the episodic memory, there must be a modification of sequence of behaviors. Anomaly is detected when the pre-conditions of behaviors from the episodic memory do not match with current environmental situations, and behavior planner re-plans to seek for a new plan to finish the given task. The score for behavior selection in tree search is calculated by

$$f = g + h, \tag{6}$$

with

$$h = h_0 + (k_1 \times c_m + K_2 \times (\frac{1}{d})) \times (1 - w), \tag{7}$$

where

$$w = min(\tfrac{n}{N}, 1),$$
$$c_m = \frac{Ranking(cogency(b_m))}{M}$$

.

Note that traditional A* is added by the cogency term and the BD term. The behavior with lower value is more likely to be selected. Pseudo-code is shown in Table. 1.

Algorithm 1 Algorithm for episodic memory augmented Cogent confabulation behavior planner

Procedure: Confabulation-based Behavior Planner(start, goal)

```
Closed_Set = {}; Open_Set = {}
previous_state[start] = empty
g_score[start] = 0; c_score[start] = 0
f_score[start] = g_score[start] +
                 heuristic_cost_estimate(start, goal)
while OpenSet is not empty do
  current_state = the node in Open_Set having
                  the lowest f_score[] value
  if current_state = goal then
    return reconstruct_behavior_sequence(previous_state,
    current_state)
  end if
  Open_Set.Remove(current_state)
  Closed_Set.Add(current_state)
  for each behavior and next_state from current_state
  do
    Open_Set.Add(next_state)
    b[next_state] = behavior
    previous_state[next_state] = current_state
    g_score[next_state]=g_score[current_state] +
      behavior_cost_between(current_state, next_state)
    h_score[next_state] =
           heuristic_cost_estimate(next_state, goal)
    c_score[next_state] = confabulation_cogency
           _ranking_score(current_state, behavior)
    d_score[next_state] = get_BD_from_Deep-ART_
    memory(ART_memory, previous_behavior, behav-
    ior)
    f_score[next_state]=g_score[next_state] + h_score
    next_state+ c_score[next_state] + d_score[next
    _state]
  end for
end while
return failure
```

end procedure

B. Learning of episodic memory augmented cogent confabulation planner

The reason of having various kinds of memory with task planner for the robot is to provide possible pool of behaviors by semantic memory and to train the planner to certain situation by episodic memory. Actually, only by using cogent confabulation behavior planner with semantic memory, training the planner is possible. However, it is biased to the training scenario much that testing it on different scenario returns poor result. Even though the result is intuitively understandable it would be better with more effective performance. Eq. (8) is the learning mechanism for cogent confabulation term (c_m) in Eq. (6). Whenever a pair of quantified situated object and quantified situated behavior is selected, its probability to be selected increases by λ and gets normalized. The f score for behavior selection decreases as the probability gets larger, since the cogency ranking is used for calculation (Eq. (3)). The reinforcement learning method is as follows:

$$\textit{Reinforce}: \text{p'}(<o_i, S_j^i > |b_m) \leftarrow p(< o_i, S_j^i > |b_m) + \lambda$$

$$\textit{Normalization}: \text{p}(<o_i, S_j^i > |b_m) \leftarrow \frac{p'(<o_i, S_j^i >|b_m)}{\sum_{n=1}^{K^{ij}} p'(<o_i, S_j^i >|b_m)} \quad (8)$$

The learning mechanism for episodic memory augmented cogent confabulation planner is as follows:

$$d_{ki} \leftarrow \lambda_d \times d_{ki}, \quad (9)$$

where d_{ki} is the BD value between two successive i^{th} and k^{th} behaviors as a pair, λ_d is the learning rate, which is multiplied whenever the i^{th} behavior following the k^{th} behavior gets observed.

V. Experiment

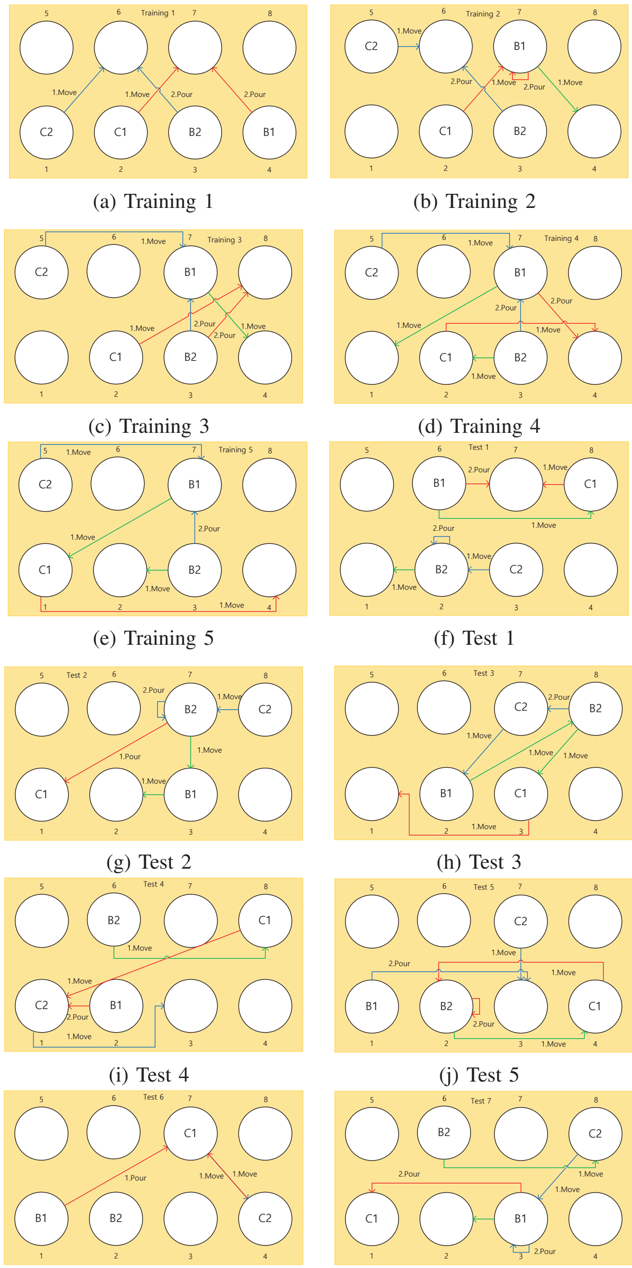
A. Experiment setting

A set of simulations to train the proposed behavior planner with episodic memory was performed. As shown in Fig. 8, five different scenarios for beverage serving was simulated to train and seven different scenarios for beverage serving to test and compare different tree search algorithms: A*, A* integrated with cogency and behavioral-causal distance value (C/BD), A* with cogency value only (C), and A* with behavioral-causal distance value (BD). Training 1 5 in Fig. 8 were run using A* only to train cogency (semantic memory) and BD (episodic memory) values which were tested separately and altogether.

This experiment aimed to check the effectiveness of episodic memory augmented to behavior planner.

B. Experimental result

Fig. 9 showed the result of experiment. Fig. 9 (a) is for the training set and (b) is for the test set. In both cases, the behavior planner with both episodic and semantic memory (A*+C/BD) traverses the smallest number of nodes. A* algorithm does not have learning mechanism, which keeps searching from being biased to certain scenarios. The planner with only cogency value (semantic memory) which trained the relationship between states and behaviors returned poor result and the planner with only BD (episodic memory) training the relationship between two successive behaviors enhances performance over the former. Once the planner defined both relationships between behaviors and state-behavior (A*+C/BD), the performance became the best



(a) Training 1 (b) Training 2 (c) Training 3 (d) Training 4 (e) Training 5 (f) Test 1 (g) Test 2 (h) Test 3 (i) Test 4 (j) Test 5 (k) Test 6 (l) Test 7

Figure 8: Plans for beverage serving task. (a)∼(e) Training scenarios. (f)∼(l) Test scenarios.

on both training and test sets in great scale. This result showed the effectiveness of utilizing the episodic memory which defines causal distance between behaviors.

VI. Discussion

The planner defining the relationship between states and behaviors (A*+C) decreased performance because it was trained and biased. However, its performance was improved when the temporal causal relationship between situated

<Training>: number of nodes planner searched through (abstract_level/manipulation_level/primitive_level)

Scenario	A*	A* + C/BD	A* + C	A* + BD
Training 1	330 / 276 / 4442	88 / 155 / 780	143 / 277 / 4448	331 / 190 / 3590
Training 2	924 / 317 / 4716	924 / 194 / 858	5705 / 314 / 4722	2071 / 208 / 3664
Training 3	1145 / 347 / 4715	778 / 221 / 986	5317 / 344 / 4721	2260 / 217 / 3704
Training 4	2834 / 337 / 4990	2844 / 212 / 1557	7728 / 340 / 4996	2947 / 241 / 4015
Training 5	464 / 238 / 3042	455 / 145 / 1043	4345 / 238 / 3045	1385 / 190 / 1988

(a) Number of nodes traversed by algorithms on training scenarios

<Testing>: number of nodes planner searched through (abstract_level/manipulation_level/primitive_level)

Scenario	A*	A* + C/BC	A* + C	A* + BD
Testing 1	918 / 307 / 4716	575/207/1660	3163/307/4722	2069/239/3885
Testing 2	557/359/4716	213/208/3638	2406/359/4722	1430/237/4203
Testing 3	201/240/3042	201/168/1409	1804/240/3045	839/182/2318
Testing 4	123/172/2768	85/172/1820	1610/172/2770	761/161/2364
Testing 5	4622/324/4990	8404/225/2482	15211/324/4996	9772/255/4083
Testing 6	748/242/2767	427/187/764	531/187/2770	748/142/2038
Testing 7	1097/310/4989	780/210/3278	5021/239/4995	1970/310/4147

(b) Number of nodes traversed by algorithms on testing scenarios

Figure 9: Number of nodes traversed by four types of planner: A*, A* with cogency and BD values, A* with cogency only, A* with BD value only. (a) Training scenarios. (b) Test scenarios.

behaviors was introduced (A*+BD). This proves that associating between situated behaviors are more effective than associating situated objects with situated behaviors; it means that the episodic memory considering temporal relationship between behaviors takes important role in planning. It is found that in planning it is difficult to confirm if the planner is trained well or not, because only with difference in initial and goal states, we cannot tell whether certain tasks are in the same category or not. Because the actual space domain is different from the planning solution domain for the robot. For instance, in the two scenarios of Fig. 8 (a) and (f), even though they look very different in space domain, considering their solutions for robot, the two successive situated behaviors 'move cup1 to location7 $\rightarrow$ pour Bottle1 in cup1' are required in both of the two scenarios. This feature affects performance of planners in the scenarios which do not look closely related to training scenarios. Categorizing tasks in terms of task planner is beyond the scope of this paper and should be studied further.

VII. Conclusion

This paper proposed a novel algorithm to integrate episodic memory with behavior planner which has semantic memory of situated affordance. A new metric BD, calculating causal distance between situated behaviors, was introduced to define temporal relationship between situated behaviors returned by episodic memory. The result showed great improvement in the number of nodes traversed to find the solution to prove that temporal relationship between situated behaviors takes important role, and consequently, the importance of episodic memory in planning. The further research is required on how to categorize and define environmental situations or status in terms of behavior planning before acquiring solution by running planner.

<video link: http://rit.kaist.ac.kr/home/ART_based_planner>

Acknowledgment

This work was supported by the Technology Innovation Program, 10045252, Development of robot task intelligence technology, supported by the Ministry of Trade, Industry, and Energy (MOTIE), Korea.

References

[1] G. A. Carpenter and S. Grossberg, "Adaptive resonance theory."

[2] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy art: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural networks*, vol. 4, no. 6, pp. 759–771, 1991.

[3] W. Wang, B. Subagdja, A.-H. Tan, and J. A. Starzyk, "Neural modeling of episodic memory: Encoding, retrieval, and forgetting," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 23, no. 10, pp. 1574–1586, 2012.

[4] J. Hawkins, S. Ahmad, and D. Dubinsky, "Hierarchical temporal memory including htm cortical learning algorithms," *Techical report, Numenta, Inc, Palto Alto http://www. numenta. com/htmoverview/education/HTM_CorticalLearningAlgorithms. pdf*, 2010.

[5] A. Al-Moadhen, R. Qiu, M. Packianather, Z. Ji, and R. Setchi, "Integrating robot task planner with common-sense knowledge base to improve the efficiency of planning," *Procedia Computer Science*, vol. 22, pp. 211–220, 2013.

[6] J.-H. Kim, S.-H. Choi, I.-W. Park, and S. A. Zaheer, "Intelligence technology for robots that think [application notes]," *Computational Intelligence Magazine, IEEE*, vol. 8, no. 3, pp. 70–84, 2013.

[7] C. Galindo, J.-A. Fernández-Madrigal, J. González, and A. Saffiotti, "Robot task planning using semantic maps," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 955–966, 2008.

[8] R. E. Fikes and N. J. Nilsson, "Strips: A new approach to the application of theorem proving to problem solving," *Artificial intelligence*, vol. 2, no. 3-4, pp. 189–208, 1971.

[9] R. Bogacz, S. M. McClure, J. Li, J. D. Cohen, and P. R. Montague, "Short-term memory traces for action bias in human reinforcement learning," *Brain research*, vol. 1153, pp. 111–121, 2007.

[10] S.-H. Cho, Y.-H. Kim, I.-W. Park, and J.-H. Kim, "Behavior selection and memory-based learning for artificial creature using two-layered confabulation," in *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*. IEEE, 2007, pp. 992–997.

[11] S.-H. Cho, S.-H. Baek, D.-H. Kim, Y.-H. Yoo, S. Cho, and J.-H. Kim, "Cogent confabulation-based hierarchical behavior planner for task performance," in *SYSTEMS, MAN, AND CYBERNETICS 2016. IEEE International Conference on*. IEEE, 2016.

[12] M. Kammer, T. Schack, M. Tscherepanow, and Y. Nagai, "From affordances to situated affordances in robotics-why context is important," in *Frontiers in Computational Neuroscience (Conference Abstract: IEEE ICDL-EPIROB 2011)*, vol. 5, no. 30, 2011.