# Context Preference-based Deep Adaptive Resonance Theory: Integrating User Preferences into Episodic Memory Encoding and Retrieval

Dick Sigmund, Gyeong-Moon Park, and Jong-Hwan Kim
School of Electrical Engineering
KAIST
Daejeon 34141, Republic of Korea
Email: {dick, gmpark, johkim}@rit.kaist.ac.kr

***Abstract*—Episodic memory which can store and recall episodes has been modeled by various research. Those models focus on encoding and retrieving the same sequence of events of episodes. In this paper, we propose context preference-based deep adaptive resonance theory (CPD-ART). CPD-ART uses a new approach in encoding and retrieving a temporal sequence of events considering subjects, preference criteria such as weather, and object contexts such as beverage. A new layer, context preference field, is added to the encoding and retrieval processes for decision making. Context preference field encodes and stores the knowledge of criteria and object contexts, along with their relations in probability weight vectors. Simulation results demonstrate that CPD-ART is able to conduct decision making analysis and retrieve the sequence of events of an episode correctly through decision making analysis based on subjects, preference criteria, and the object contexts.**

## I. Introduction

In psychology and neuroscience, human long-term memory is generally divided into two categories: implicit memory and explicit memory [1], [2]. Implicit memory works under human unconsciousness while explicit memory works under human consciousness [2]. The most common form of implicit memory is procedural memory [3] which helps humans to perform tasks such as riding a bicycle, writing with a pencil, etc. On the other hand, explicit memory consists of two divisions in general: semantic memory and episodic memory [4]. Semantic memory refers to the memory that stores factual information accepted between each other [5] such as knowing what a phone is, South Korea capital is Seoul, etc. Episodic memory refers to memory that stores and retrieves past events and personal experiences; it stores those events in time sequential form [6], [7]. Some examples of episodic memory are remembering our graduation day, our first day of school, etc.

A lot of research in modeling episodic memory has been conducted in the past few years. REM-II [8] modeled episodic memory based on Bayesian features. A. Nuxoll et al. [9] also defined a design space for episodic memory model and implemented it within a cognitive architecture. However, these models can only store simple events and their basic relations. W. Wang et al. [10] developed a novel episodic memory model based on fusion adaptive resonance theory (fusion ART) [11]. This model can encode complex events and recall them even with partial and erroneous cues. Later, forgetting mechanism was also included in [10] and proposed as episodic memory-adaptive resonance theory (EM-ART) [12].

Applications of episodic memory based on EM-ART are used in several different fields: human daily activity pattern [13], autonomous robot [14]–[17], etc. Nevertheless, EM-ART is found out to fail in storing duplicate events in an episode [18] and also retrieval error may occur depending on the number of events in stored episodes [16]. In order to solve these problems, Deep ART [19] which also works based on fusion ART neural model was proposed. Deep ART uses a different kind of encoding and decoding methods in the event field such that it can store duplicate events as well. Also, complement coding is removed during the retrieval process to solve the retrieval error due to the different number of events in stored episode. Thus, Deep ART is robust to both of encoding and retrieval errors. By this time, research about episodic memory focused on storing and recalling the episodes; research about encoding and retrieval based on user preferences has not deeply studied. G. Sieber et al. [20] proposed a resource description framework (RDF)-based episodic memory component that takes account of user preferences from past interactions. Nonetheless, this model is limited only to language processing since it was applied to an artificial companion for having a dialogue.

In this paper, we propose a context preference-based Deep ART (CPD-ART) which integrates user preferences into episodic memory model. CPD-ART retrieves not only stored event's input vectors and attributes, but also different kinds of inputs depending on the subjects and preference criteria. Adapting Deep ART procedure, CPD-ART encodes all the episodes, subjects, preference criteria, and object contexts. In the retrieval process, context preference field is introduced. Context preference field acts as a decision maker in which subjects, preference criteria and object context field of a particular input channel are considered to retrieve event's input vectors and attributes. Beside determining the decision of event's input vectors and attributes, context preference field also learns and updates the relations between those criteria and

978-1-5090-6182-2/17/$31.00 ©2017 IEEE 

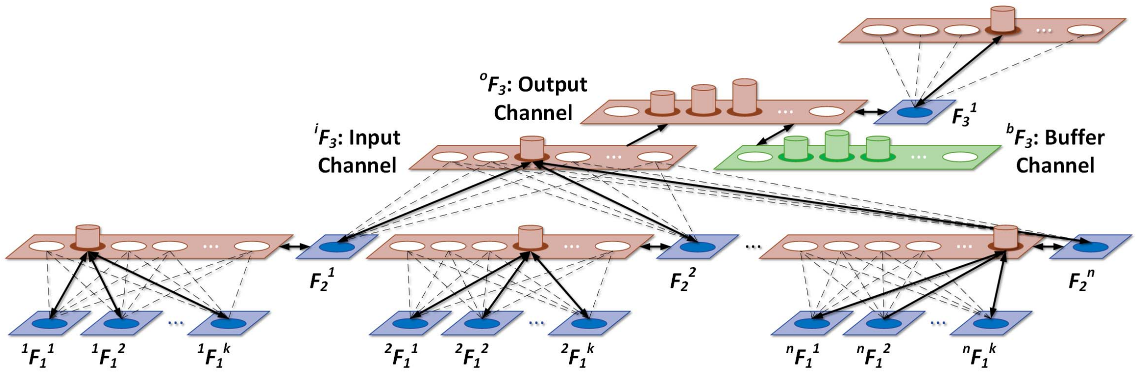



Fig. 1: Deep ART architecture with four layers structure: attribute field, input field, event field, and episode field.

object contexts after the retrieval process. The effectiveness of the proposed CPD-ART is demonstrated through computer simulations for the task intelligence of a robot in which it learns some tasks and retrieve them based on provided criteria.

This paper is organized as follows. Section II presents an overview of Deep ART. Section III describes the procedure of the proposed CPD-ART. Simulation results along with the discussion regarding the effectiveness of CPD-ART are presented in Section IV. Finally, concluding remarks follow in Section V.

## II. Deep ART Architecture

Deep ART is an unsupervised learning neural model that can learn and retrieve episodes of temporal sequences of events as in episodic memory. It consists of several layers depending on the usage, but in general four layers are used for an episodic memory: attribute field, input field, event field, and episodic field. The architecture of Deep ART is shown in Fig. 1.

### A. Input and Event Encoding

Deep ART receives input vectors ${}^n\mathbf{I}^k$ from attribute channels ${}^nF_1^k$ which contains features of input channels in the input field $F_2$. The process to encode input consists of some basic operations of fusion ART [11] as follows:

*1) Complement Coding:* Input vectors ${}^n\mathbf{I}^k$ that are passed to attribute channel ${}^nF_1^k$ go through complement coding process as follows:

$$\begin{aligned} {}^n\bar{\mathbf{I}}^k &= 1 - {}^n\mathbf{I}^k, \\ {}^n\mathbf{x}^k &= \{{}^n\mathbf{I}^k, {}^n\bar{\mathbf{I}}^k\}. \end{aligned} \tag{1}$$

*2) Code activation:* The input vectors ${}^n\mathbf{x}^k$ are activated at the $j^{th}$ node of input field $F_2$ by the following choice function:

$${}^n\mathbf{T}_j = \sum_{k=1}^{j} {}^n\gamma^j \frac{|{}^n\mathbf{x}^k \wedge^n \mathbf{w}_j^k|}{{}^n\alpha^j + |{}^n\mathbf{w}_j^k|}. \tag{2}$$

*3) Code competition:* All nodes in the $n^{th}$ input channels go through a competition process. The winner, indexed at $J$, is selected by checking which node has the highest choice function value as follows:

$${}^n\mathbf{T}_J = max\{{}^n\mathbf{T}_j : \text{ for all } F_2 \text{ node } j\}. \tag{3}$$

*4) Template matching:* This process checks whether resonance occurs at $J^{th}$ node. Resonance occurs if the following match function ${}^nm_J^k$ is larger than the vigilance parameter ${}^n\rho^k$:

$${}^nm_J^k = \frac{|{}^n\mathbf{x}^k \wedge^n \mathbf{w}_j^k|}{|{}^n\mathbf{w}_j^k|} \geq {}^n\rho^k. \tag{4}$$

*5) Template learning:* If resonance occurs, the weight for every channel $k$ will be updated as follows:

$${}^nw_J^{k(new)} = (1 - {}^n\beta^k){}^nw_J^{k(old)} + {}^n\beta^k({}^nx^k \wedge {}^nw_j^{k(old)}). \tag{5}$$

After input encoding, input channels of input field $F_2^n$, which specify a particular event, are encoded in the input channel of event field ${}^iF_3$. Event encoding also has the same process as the input encoding in which input field $F_2$ is the input and the input channel of event field ${}^iF_3$ is the output of the fusion ART.

### B. Episode Encoding

Episode field $F_4$ encodes the sequences of events provided by the event field $F_3$ as shown in Fig. 2. The encoding process of a temporal sequence of events is implemented using the following equations:

$$\begin{aligned} {}^b\mathbf{x}_n &= {}^ow\ {}^o\mathbf{y}_{n-1} \\ {}^o\mathbf{y}_n &= {}^iw\ {}^i\mathbf{x}_n + {}^bw\ {}^b\mathbf{x}_n \\ &= {}^iw\ {}^i\mathbf{x}_n + {}^bw\ {}^ow\ {}^o\mathbf{y}_{n-1} \\ &= {}^iw \textstyle\sum_{k=0}^{n-1} ({}^bw\ {}^ow)^k\ {}^i\mathbf{x}_{n-k} \end{aligned} \tag{6}$$

where ${}^i\mathbf{x}_n$ and ${}^b\mathbf{x}_n$ are the input and buffer vectors, ${}^o\mathbf{y}_n$ is the output vector, and ${}^iw$, ${}^bw$, ${}^ow$ are the input, buffer, output

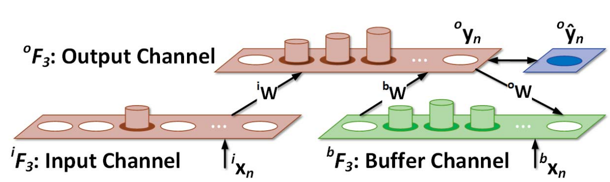



Fig. 2: Encoding process of event sequences in Deep ART.

weight vectors, respectively. Note that ${}^b\mathbf{x}_1$ is initialized with zero which means that ${}^o\mathbf{y}_0$ is a zero vector at the beginning.

After the encoding process using (6), the output vectors ${}^o\mathbf{y}_n$ are normalized before inputted to the episode field $F_4$. The normalized output ${}^o\hat{\mathbf{y}}_n$ is simply calculated by dividing the output vectors with the maximum positional number of the output vectors.

Finally, episode in episode field is encoded by using fusion ART, the same method described in subsection A. The normalized output vectors of event field, ${}^o\hat{\mathbf{y}}_n$, act as the input of fusion ART and the encoded episode in episode field acts as the output of fusion ART.

### C. Episode, Event, and Input Retrieval

After learning phase, Deep ART can retrieve a specific episode along with its temporal sequence of events and input vectors. The summary of retrieval process is described below.

*1) Episode retrieval:* The chosen episode, denoted as $J^{th}$ node, is retrieved by reading out the corresponding weight vectors which is the normalized output vectors in the events' sequence encoding.

$$ {}^o\hat{\mathbf{y}}_n = {}^J w_3. \tag{7}$$

*2) Sequence of events retrieval:* This process is basically the reverse of the encoding process in which at first the normalized output vector ${}^o\hat{\mathbf{y}}_n$ is denormalized by multiplying it with the maximum positional number. Then, the denormalized output vectors ${}^o\mathbf{y}_n$ is inputted to the buffer channel which later yields one element to the input channels of event field one at a time.

*3) Event and input retrieval:* Same as episode retrieval, the chosen node in event field is retrieved by reading out the corresponding weights between event field and input field as shown in (7), yielding input channels of the input field. This process is also performed in the input field to retrieve input attributes.

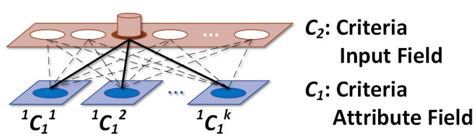
## III. CPD-ART Architecture

Same as Deep ART, CPD-ART encodes and retrieves the temporal sequences of events. However, CPD-ART also considers subject and preference criteria knowledge during the encoding and retrieval processes. The main difference between proposed model and Deep ART appears in the type of data or fields to be encoded and also the procedure of the retrieval process which will be presented in this section.

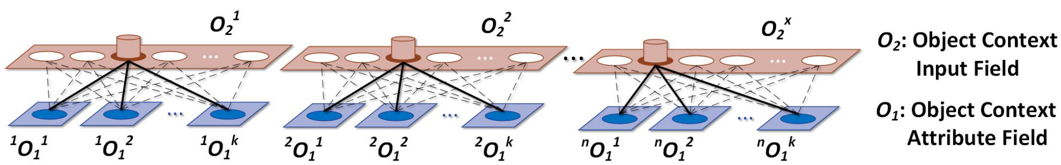
### A. Encoding Process

In Deep ART, encoding process is carried out only for the attributes, inputs, events, and episodes. In CPD-ART, subjects, preference criteria, and object contexts are also encoded. Subject is the person that CPD-ART interacts with; preference criteria are the factors that CPD-ART needs to consider and evaluate in order to make decision, such as weather, place, time, etc. The decision itself is defined as object context, which is the description about an object along with its detailed characteristics, e.g. water with hot temperature. There can be several object contexts depending on the episode and event the memory is dealing with.

Fig. 3 shows the additional encoding process in CPD-ART. Both subjects and preference criteria are encoded with fusion ART as shown in Fig. 3a. The encoding process consists of two layers: criteria attribute field $C_1$ and criteria input field $C_2$. The criteria attribute field $C_1$ has several channels ${}^1C_1^k$ which could be a subject channel and some preference criteria channels. Similarly, object context is also encoded with fusion ART as shown in Fig. 3b. It is also consisted of two layers: object context attribute field $O_1$ and object context input field $O_2$. The object context attribute field $O_1$ has several channels ${}^nO_1^k$ that resemble the characteristics of the corresponding object context input channel $O_2^x$. Several object context input channels $O_2^x$ can be created depending on the number of object contexts. After the encoding process, the criteria input field $C_2$




(a)




(b)

Fig. 3: (a) Criteria encoding process in CPD-ART in which the criteria attribute field $C_1$ has several channels ${}^1C_1^k$ that represents a subject channel and some preference criteria channels. (b) Object context encoding process in CPD-ART in which there are several object context input channels $O_2^x$ depending on the numbers of the object contexts; and also the object context attribute field $O_1$ has several channels ${}^nO_1^k$ that represents a particular object context and its characteristics.

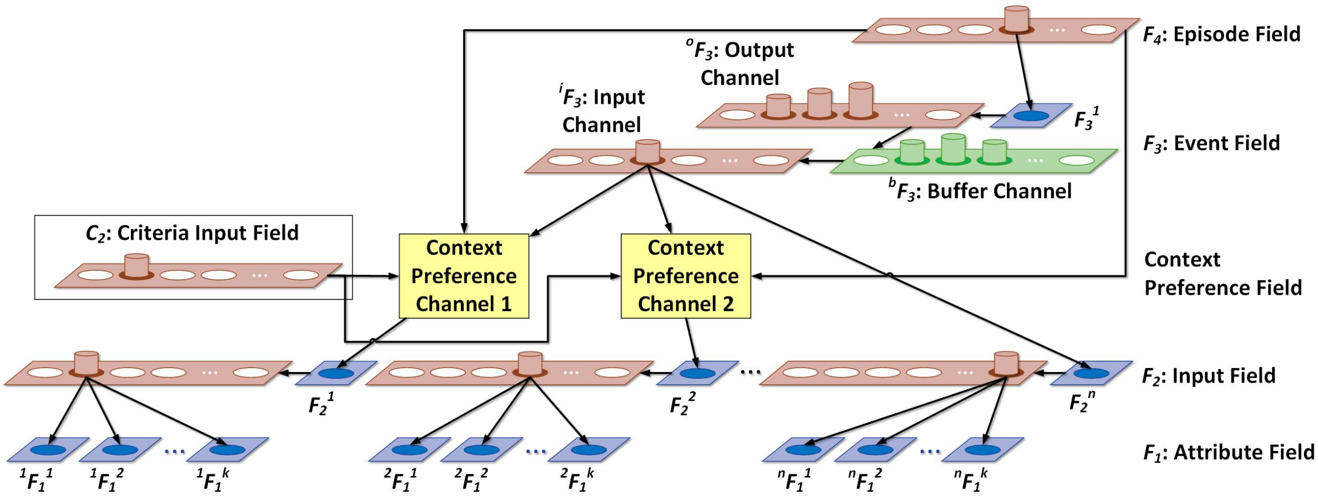



Fig. 4: Episode retrieval process in CPD-ART. This figure shows the addition of context preference channels to the first and second channel of input field, $F_2^1$ and $F_2^2$, which means that these two channels require decision making analysis on the retrieval process. On the other hand, the context preference channel is not connected to the $n^{th}$ channel $F_2^n$ because it does not require decision making analysis.

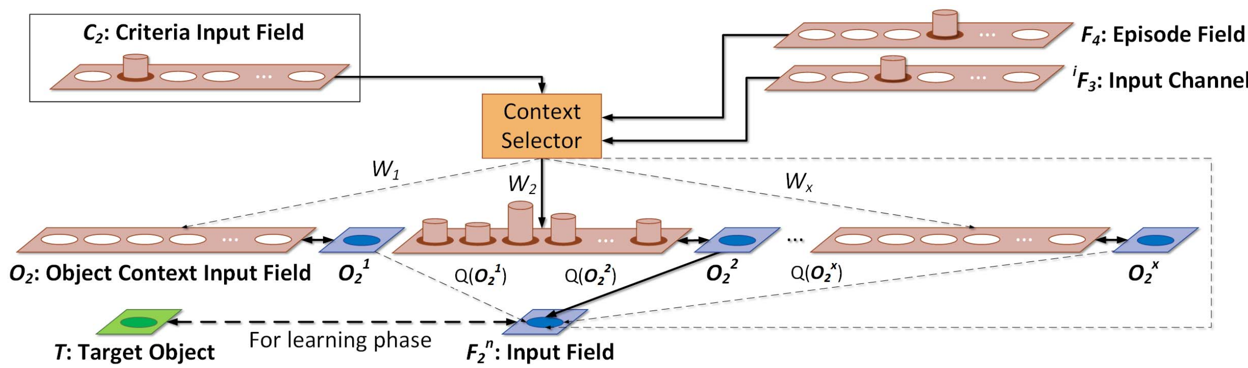



Fig. 5: Details of a context preference channel in CPD-ART. It consists of criteria input field $C_2$, episode field $F_4$, and input channel of event field ${}^iF_3$ as inputs; several object context input channels $O_2^x$ in the object context input field $O_2$; and input field $F_2^n$ as output. In this figure, the context selector chose the second object context input channel $O_2^2$ through the probability weight vector $W_2$ for decision making analysis (marked as solid black line). Target object $T$ is only used in the learning phase to update the probability weight vectors $W_x$.

and object context input field $O_2$ in Fig. 3 will be used for decision making analysis in the retrieval process.

### B. Retrieval Process

The retrieval process of the episode and sequence of events in CPD-ART follows the same procedure as in Deep ART. However, context preference field is added between event field and input field as shown in Fig. 4. Context preference field consists of one or more context preference channels. These channels are connected only to the input field channels that require decision making analysis. In Fig. 4, two context preference channels are connected to the first and the second input field channel, $F_2^1$ and $F_2^2$, but not connected to the $n^{th}$ input field channel, $F_2^n$. This means that the first and the second input field channel need decision making analysis in retrieval process, but the $n^{th}$ input field channel does not need it.

Fig. 5 illustrates the detailed architecture of a context preference channel. This channel is constructed with several object context input channels $O_2^x$ from the object context input field $O_2$ encoded previously using Fig. 3b. Furthermore, context preference channel receives three inputs: episode field $F_4$, event input channel ${}^iF_3$, and criteria input field $C_2$. Episode field and event input channel play roles in providing knowledge about current episode and event that CPD-ART is dealing with. On the other hand, criteria input field plays a role as the factor that CPD-ART need to evaluate for making decision in the current episode and event.

Context selector in Fig. 5 decides where the criteria input field $C_2$ needs to be forwarded considering the episode field $F_4$ and the event input channel ${}^iF_3$. The connection between context selector and each object context input channel $O_2^x$ are the probability weight vectors $W_x$ where $x$ is the index of corresponding object context input channel $O_2^x$. This weight vector functions as the main component in analyzing decision making. Since probability weight vector resembles the relationship between criteria and object context, it has a size of $N \times M$ for each object context input channel $O_2^x$ where $N$ is the number of nodes in criteria input field $C_2$ and $M$ is the number of nodes in the corresponding object context input channel $O_2^x$. Each probability weight vector $W_x$ is initialized by the following equation:

$$W_x = \left\{ \frac{1}{M} \right\} \text{ for all elements.} \tag{8}$$

The context preference channel works in either two phases: prediction phase or learning phase. Prediction phase is selected if a target object $T$ is not provided. On the other hand, learning phase is selected if a target object $T$ is provided. The target object is the desired decision in input field $F_2^n$, provided by the subject.

*1) Prediction Phase:* The process of this phase begins from the context selector which determines whether the corresponding episode and event input channel require decision making analysis. If it does not require this analysis, the event input channel ${}^iF_3$ is simply passed to the corresponding input field as follows:

$$F_2^n = {}^iF_3. \tag{9}$$

However, if it requires decision making analysis, context selector will connect criteria input field $C_2$ to a selected probability weight vector. The selected weight vector $W_x$ is chosen by considering the event input channel ${}^iF_3$ and the episode field $F_4$. For instance, if the current episode is episode 2 and the event input channel is correlated with the second object context input channel $O_2^2$, then the selected connection is the one with probability weight vector $W_2$. After selecting the proper probability weight vector, the decision making is analyzed by calculating the probability relation between the criteria and the corresponding object context input channel $x$ as follows:

$$O_2^x = C_2 W_x. \tag{10}$$

Afterwards, a winner-take-all strategy, denoted as $Q(.)$, in (11) is conducted to finalize the decision by assigning the value of each node in the object context input channel $O_2^x$ as 0 or 1 as follows:

$$Q_y(O_2^x) = \begin{cases} 1, & \text{for } {}^yO_2^x = max(O_2^x) \\ 0, & \text{for others} \end{cases} \tag{11}$$

where y is the index of the node in the corresponding object context input channel $O_2^x$. Finally, the output vector $Q(O_2^x)$ is forwarded to the input field $F_2^n$ as the final decision made by the context preference channel. Then from the input field $F_2^n$, the retrieval process continues to the input and attribute retrieval as in Deep ART.

*2) Learning Phase:* Learning phase is selected if an target object $T$ is given to CPD-ART. First, the target object received from subject is converted into binary representation so that the target object matches with the object context input channel representation. Then the target object is forwarded to the input field $F_2^n$ for further retrieval process. After retrieval process is over, update on CPD-ART knowledge is carried out by first predicting the decision based on the current probability weight vector using (8) to (11). Finally, the weight vector is updated by comparing the predicted decision $Q(O_2^x)$ and the given target object $T$ as follows:

$$W_x^{z(new)} = W_x^{z(old)} + \eta(T - Q(O_2^x)) \text{ , } z \in N \tag{12}$$

where $z$ is the index of the activated node in the criteria input field, $\eta$ is the learning rate, and $T$ is the target object in binary representation.

## IV. Simulation

We evaluated the performance of CPD-ART by simulating two tasks of a service robot. More details about these two tasks and their procedure are shown in Table I.

### A. Simulation Environment

In the simulation, we used two input channels in the input field layer: action and object. There is no context preference channel connected to the action input channel, but there is one connected to the object input channel. This context preference channel has two object context input channels: beverage and pizza. The beverage context input channel have two types of attributes: 1 (coffee, tea, or water) and 2 (hot or iced). The pizza context input channel also have two types of attributes: 1 (ham pizza or seafood pizza) and 2 (thin or thick).

The criteria defined for context preference channel are subjects, time, and weather. Subjects have John and Mary;

TABLE I: Two Task Episodes Used in Simulation

| **Serving Beverage** | | **Serving Pizza** | |
|---|---|---|---|
| *Action* | *Object* | *Action* | *Object* |
| Grasp | Cup | Grasp | Pizza |
| MoveTo | Vending | PutOn | Plate |
| Release | Cup | Grasp | Plate |
| Press | Beverage | MoveTo | Microwave |
| Grasp | Cup | Open | Microwave |
| MoveTo | Table | PutIn | Plate |
| Release | Cup | Close | Microwave |
| | | Press | MicrowaveButton |
| | | Open | Microwave |
| | | Grasp | Plate |
| | | Close | Microwave |
| | | MoveTo | Table |
| | | Release | Plate |

TABLE II: First Scenario of Episode Retrieval during A Week

| No. | Subject | Preference Criteria | | Episode to Retrieve | Target Object | No. of Retrievals |
|---|---|---|---|---|---|---|
| | | Time | Weather | | | |
| 1 | John | Morning | Warm | Beverage | Hot Water | 2 |
| | | | | Pizza | Thick Seafood | 3 |
| 2 | John | Noon | Hot | Beverage | Iced Coffee | 3 |
| | | | | Pizza | Thick Ham | 2 |
| | | | | | Thin Seafood | 1 |
| 3 | John | Night | Cold | Beverage | Hot Water | 1 |
| | | | | | Hot Tea | 2 |
| | | | | Pizza | Thin Seafood | 2 |
| 4 | Mary | Morning | Warm | Beverage | Hot Tea | 3 |
| | | | | Pizza | Thin Ham | 2 |
| 5 | Mary | Noon | Hot | Beverage | Iced Water | 2 |
| | | | | Pizza | Thin Seafood | 2 |
| 6 | Mary | Night | Cold | Beverage | Hot Water | 2 |
| | | | | Pizza | Thin Seafood | 1 |
| | | | | | Thick Ham | 2 |

TABLE III: Second Scenario of Episode Retrieval without Providing any Target Object

| No. | Subject | Preference Criteria | | Episode to Retrieve | Retrieval Results |
|---|---|---|---|---|---|
| | | Time | Weather | | |
| 1 | John | Morning | Warm | Beverage | Hot Water |
| | | | | Pizza | Thick Seafood |
| 2 | John | Noon | Hot | Beverage | Iced Coffee |
| | | | | Pizza | Thick Ham |
| 3 | John | Night | Cold | Beverage | Hot Tea |
| | | | | Pizza | Thin Seafood |
| 4 | Mary | Morning | Warm | Beverage | Hot Tea |
| | | | | Pizza | Thin Ham |
| 5 | Mary | Noon | Hot | Beverage | Iced Water |
| | | | | Pizza | Thin Seafood |
| 6 | Mary | Night | Cold | Beverage | Hot Water |
| | | | | Pizza | Thick Ham |

time criterion has morning, noon, and night; weather criterion has cold, warm, and hot. The learning rate $\eta$ used in the context preference channel learning phase was 0.1.

### B. Simulation Scenarios

At first, the two episodes, object contexts, subject and preference criteria are encoded in CPD-ART. After encoding process, retrieval process is carried out using two scenarios as shown in Table II and III.

The first scenario (Table II) is retrieving episode with subject, preference criteria, and provided target object. In this scenario, CPD-ART works in learning phase at the context preference channel. Target object in Table II is the order or object that subject defined for the robot. For instance, event sequence No. 2 means John is ordering iced coffee to the robot at noon when it is hot. In this event sequence, 'John' is the subject, 'iced coffee' is the target object, 'noon' and 'hot' are the preference criteria. Additionally, the no. of retrievals is 3 which means that John already ordered this object three times in a week. Note that the retrieval of each event sequence in the first scenario is randomly carried out throughout the week.

TABLE IV: Details of the Two Tasks Retrieved by CPD-ART from Second Scenario when John Ordered Beverage and Pizza at A Warm Morning

| Serving Beverage | | Serving Pizza | |
|---|---|---|---|
| *Action* | *Object* | *Action* | *Object* |
| Grasp | Cup | Grasp | Thick Seafood Pizza |
| MoveTo | Vending | PutOn | Plate |
| Release | Cup | Grasp | Plate |
| Press | Hot Water | MoveTo | Microwave |
| Grasp | Cup | Open | Microwave |
| MoveTo | Table | PutIn | Plate |
| Release | Cup | Close | Microwave |
| | | Press | MicrowaveButton |
| | | Open | Microwave |
| | | Grasp | Plate |
| | | Close | Microwave |
| | | MoveTo | Table |
| | | Release | Plate |

On the contrary, the second scenario (Table III) is retrieving episode without providing any target object. In this scenario, CPD-ART works in prediction phase at the context preference channel. In this case, John does not give an order to the robot at noon when it is hot. However, the robot is expected to retrieve a particular beverage or pizza based on its knowledge.

### C. Simulation Results

Fig. 6 and 7 show the initial probability weight vectors of beverage context input channel and pizza context input channel, respectively, in accordance with the subject and

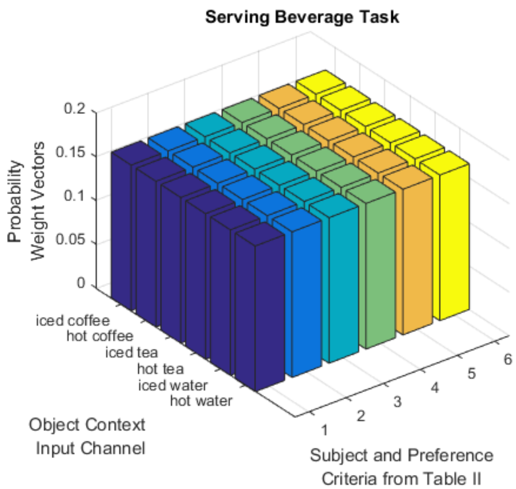



Fig. 6: Initial probability weight vectors plot for serving beverage.

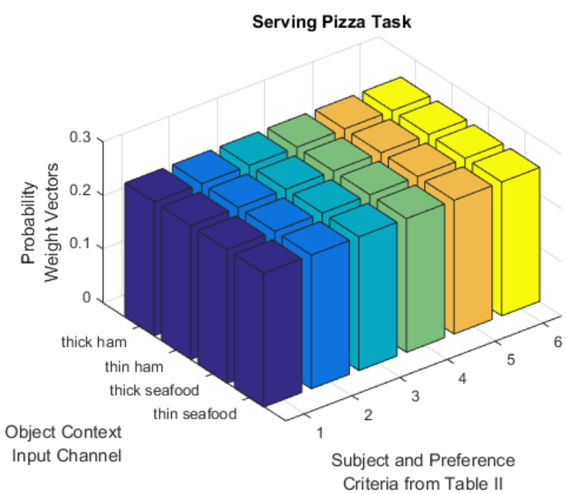



Fig. 7: Initial probability weight vectors plot for serving pizza.

preference criteria. We can see that for beverage and pizza context input channels, the values of the probability weight vectors are 0.167 and 0.25 which follows (8). Note that the subject and preference criteria numbers in the figures are based on the event sequence numbers in Table II. In other words, no. 1 corresponds to John, morning, and warm; no. 2 corresponds to John, noon, and hot; and so on.

After running the first scenario in Table II for retrieval process, the probability weight vectors of both beverage and pizza context input channels are updated as in Fig. 8 and 9, respectively. From these results, we can imply that if John orders beverage and food at noon when it is hot, CPD-ART would probably make decision to retrieve 'iced coffee' and 'thick ham pizza' since they have the highest probability value in each object context input channel. On the other hand, if Mary orders beverage and food at noon when it is hot, CPD-ART would probably make decision to retrieve

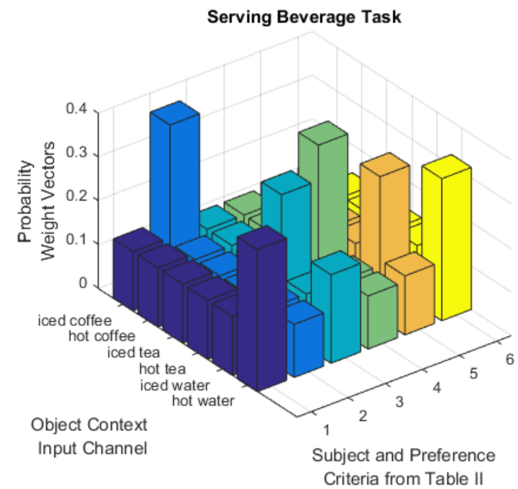



Fig. 8: Probability weight vectors plot for serving beverage after first scenario.

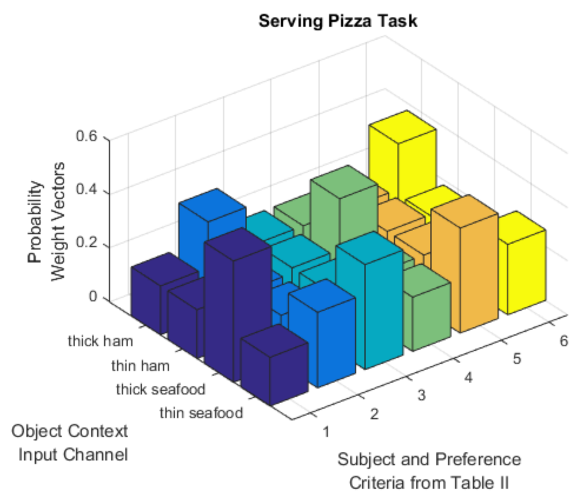



Fig. 9: Probability weight vectors plot for serving pizza after first scenario.

'iced water' and 'thin seafood pizza'. Thus, learning phase of the context preference field was successfully performed as CPD-ART could learn the preference of the subject and retrieve them accordingly.

Table III shows the retrieval results of the second scenario which works in prediction phase at the context preference field. We can see that all the results retrieved by CPD-ART decision making analysis are the objects with highest values of probabilities weight vectors in Fig. 8 and 9. These show that context preference-based retrieval process is carried out correctly according to the knowledge learned from the first scenario by context preference field.

For further analysis, episode retrieval results of 'John ordering beverage and pizza at morning when it is warm' are presented in Table IV. As shown, CPD-ART is able to retrieve duplicate events such as 'release cup' and 'grasp cup' in serving beverage task and also 'grasp plate', 'open microwave', and 'close microwave' in serving pizza task. Besides, object in certain events is also changed based on the criteria (shown as underlined text): 'hot water' in serving beverage task and 'thick seafood pizza' in serving pizza task. This implies that CPD-ART can retrieve not only episode with duplicate events but also different kinds of object based on subjects and preference criteria provided.

## V. Conclusion

This paper proposed a new approach of user preferences integration in episodic memory encoding and retrieval processes based on CPD-ART. It works under Deep ART framework with context preference field introduced in both encoding and retrieval processes. The context preference field enables CPD-ART to make decision based on subject and preference criteria. Our simulation shows that CPD-ART is capable of not only retrieving the encoded sequence of events correctly, but also retrieving different kinds of input of an event based on the encoded criteria and episode context.

For future works, we will find a method to indicate a new preference criterion or object so that CPD-ART can learn a new knowledge without any problem. In addition, we will also consider to include the effect of emotion in our memory architecture.

## Acknowledgment

This work was partly supported by the ICT R&D program of MSIP/IITP [2016-0-00563, Research on adaptive machine learning technology development for intelligent autonomous digital companion] and the Technology Innovation Program, 10045252, Development of robot task intelligence technology, supported by the MOTIE), Korea.

## References

[1] R. C. Atkinson and R. M. Shiffrin, "Human memory: A proposed system and its control processes," *Psychology of Learning and Motivation*, vol. 2, pp. 89–195, 1968.

[2] P. Graf and D. L. Schacter, "Implicit and explicit memory for new associations in normal and amnesic subjects," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 11, no. 3, p. 501, 1985.

[3] H. L. Roediger, "Implicit memory: Retention without remembering." *American Psychologist*, vol. 45, no. 9, p. 1043, 1990.

[4] E. Tulving, "Episodic and semantic memory 1," *Organization of Memory. London: Academic*, vol. 381, no. 4, pp. 382–404, 1972.

[5] K. McRae and M. N. Jones, *Semantic Memory*. The Oxford Handbook of Cognitive Psychology, 2013.

[6] E. Tulving, *Elements of episodic memory*. New York: Oxford University Press, 1983.

[7] E. Tulving, "Episodic memory: From mind to brain," *Annual Review of Psychology*, vol. 53, no. 1, pp. 1–25, 2002.

[8] S. T. Mueller and R. M. Shiffrin, "REM-II: A model of the developmental co-evolution of episodic memory and semantic knowledge," in *Proceedings International Conference on Learning and Development (ICDL)*, Bloomington, IN, USA, 2006.

[9] A. M. Nuxoll and J. E. Laird, "Extending cognitive architecture with episodic memory," *Ann Arbor*, vol. 1001, pp. 48 109–2121, 2007.

[10] W. Wang, B. Subagdja, A.-H. Tan, and J. A. Starzyk, "A self-organizing approach to episodic memory modeling," in *Proceedings IEEE International Joint Conference on Neural Networks (IJCNN)*, 2010, pp. 1–8.

[11] A.-H. Tan, G. A. Carpenter, and S. Grossberg, "Intelligence through interaction: Towards a unified theory for learning," in *Proceedings International Symposium on Neural Networks*. Springer, 2007, pp. 1094–1103.

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

[13] S. Gao and A.-H. Tan, "User daily activity pattern learning: A multi-memory modeling approach," in *Proceedings IEEE International Joint Conference on Neural Networks (IJCNN)*, 2014, pp. 1542–1548.

[14] F. Leconte, F. Ferland, and F. Michaud, "Fusion adaptive resonance theory networks used as episodic memory for an autonomous robot," in *International Conference on Artificial General Intelligence*. Springer, 2014, pp. 63–72.

[15] B.-S. Yoo, Y.-H. Yoo, W.-R. Ko, S.-J. Lee, S.-H. Cho, and J.-H. Kim, "Realization of task intelligence based on the ioa for assistive robots," in *Proceedings International Conference on Artificial Intelligence (ICAI)*, Las Vegas, NV, USA, Jul. 2015, pp. 793–798.

[16] G.-M. Park, Y.-H. Yoo, and J.-H. Kim, "REM-ART: Reward-based electromagnetic adaptive resonance theory," in *Proceedings International Conference on Artificial Intelligence (ICAI)*, Las Vegas, NV, USA, Jul. 2015, pp. 805–811.

[17] Y.-H. Yoo and J.-H. Kim, "Procedural memory learning from demonstration for task performance," in *Proceedings IEEE International Conference on Systems, Man, and Cybernetics*, Hong Kong, China, Oct. 2015, pp. 2435–2440.

[18] B. Subagdja and A.-H. Tan, "Neural modeling of sequential inferences and learning over episodic memory," *Neurocomputing*, vol. 161, pp. 229–242, 2015.

[19] G.-M. Park and J.-H. Kim, "Deep adaptive resonance theory for learning biologically inspired episodic memory," in *Proceedings IEEE International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, Jul. 2016.

[20] G. Sieber and B. Krenn, "Towards an episodic memory for companion dialogue," in *International Conference on Intelligent Virtual Agents*. Springer, 2010, pp. 322–328.