Modeling Autobiographical Memory in Human-Like Autonomous Agents

Di Wang[1], Ah-Hwee Tan[2] and Chunyan Miao[1,2]
[1]Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly (LILY) and
[2]School of Computer Engineering
Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
{wangdi,asahtan,ascymiao}@ntu.edu.sg

ABSTRACT

Although autobiographical memory is an important part of the human mind, there has been little effort on modeling autobiographical memory in autonomous agents. With the motivation of developing human-like intelligence, in this paper, we delineate our approach to enable an agent to maintain memories of its own and to wander in mind. Our model, named Autobiographical Memory-Adaptive Resonance Theory network (AM-ART), is designed to capture autobiographical memories, comprising pictorial snapshots of one's life experiences together with the associated context, namely time, location, people, activity, and emotion. In terms of both network structure and dynamics, AM-ART coincides with the autobiographical memory model established by the psychologists, which has been supported by neural imaging evidence. Specifically, the bottom-up memory search and the top-down memory readout operations of AM-ART replicate how the brain encodes and retrieves autobiographical memories. Furthermore, the wandering in reminiscence function of AM-ART mimics how human wanders in mind. For evaluations, we conducted experiments on a data set collected from the public domain to test the performance of AM-ART in response to exact, partial, and noisy memory retrieval cues. Moreover, our statistical analysis shows that AM-ART can simulate the phenomenon of wandering in reminiscence.

Categories and Subject Descriptors

I.2.0 [**ARTIFICIAL INTELLIGENCE**]: General—*Cognitive simulation*; I.2.11 [**ARTIFICIAL INTELLIGENCE**]: Distributed Artificial Intelligence—*Intelligent agents*

General Terms

Algorithms, Experimentation

Keywords

Cognitive model, Computational autobiographical memory model, Memory storage and retrieval, Wander in reminiscence

1. INTRODUCTION

Autobiographical memory is "a system that encodes, stores and guides retrieval of all episodic information related to our personal experiences" [4]. Although autobiographical memory is an important part of the human mind, there has been little effort on modeling autobiographical memory in autonomous agents. With the motivation of developing human-like intelligence, in this paper, we delineate our approach to enable an agent to encode and retrieve the autobiographical memories and to wander in reminiscence. Reminiscence, often being investigated as a function of autobiographical memory, plays a critical role in self-acceptance and self-change [4]. In this paper, we refer to *wandering in reminiscence* as recalling a sequence of contextually connected autobiographical memory across different episodes of life events. This is the typical objective of reminiscence therapies often being used with senior citizens to improve their psychological and cognitive well-being [16].

The main focus of our paper is to show how our agent can encode and retrieve autobiographical memory (**AM**) like human brains do [7], especially when dealing with noisy retrieval cues. Moreover, our agent can emulate the wandering in reminiscence effect by retrieving contextually connected memories across episodes of events. To the best of our knowledge, our model is the first research that simulates the phenomenon of wandering in autobiographical memories. Our agent is web-based, which allows users to upload their memories and even to import memories from popular online social networks, such as Facebook. A preliminary demonstration prototype of our AM agent has been described in [29]. In this paper, we present the detailed design and the dynamics of our computational AM model and show how it can be considered as human-like.

Our model, named Autobiographical Memory-Adaptive Resonance Theory network (**AM-ART**), follows the hierarchical structure of the AM model established by psychologists [7], which has been supported by neural imaging evidence [1]. More importantly, the dynamics of AM-ART coincide with the three stages of the generative memory retrieval process identified in [7]. AM-ART adopts the dynamics of a self-organizing neural network named fusion ART [24], but extends the network structure to a three-level hierarchy. The bottom layer of AM-ART encodes event-specific knowledge comprising 5W1H, namely time (when), location (where), people (who), activity (what), imagery (which), and emotion (how). The middle layer encodes events by associating the event-specific knowledge and the top layer encodes episodes by associating related events. Memory retrieval in AM-ART first takes place in the bottom layer, where the retrieval cue is presented. By following the bottom-up memory search procedure, the corresponding event and episode can be identified in the middle and top layers, respectively. Consequently, we can retrieve them by performing the top-down memory readout procedure. One of the novel features of AM-ART is its capability to emulate human mind in wandering in reminiscence. Specifically, AM-ART can iteratively perform the gener-

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

ative retrieval of autobiographical memories by mutating the retrieval cue with regulated noise in each iteration.

Experiments have been conducted, wherein AM-ART is used to capture a person's autobiographical memory based on a data set collected from the public domain. After encoding the memories, we perform memory retrieval using three types of cues, namely exact, partial, and noisy ones. To demonstrate the advantage of AM-ART, we compare its performance against the standard keyword-based query method, which is employed by many existing photo or memory repositories and the recently proposed computational AM models (e.g., [12]). The comparisons show that AM-ART performs robust and flexible memory retrievals, especially in response to noisy cues. The ability to handle uncertainty is one of the key functionalities of our human brains, which has not been well catered by many existing photo or memory repositories. Furthermore, to evaluate the performance of AM-ART in wandering in reminiscence, we devise three measurement metrics, namely coverage, diversity, and relatedness. We compare the performance of AM-ART in different parameter settings against each other and two baseline retrieval methods, namely sequential retrieval and random retrieval. Moreover, we propose two null hypotheses that mind wandering simulated by AM-ART performs on the same level as the baseline methods. The statistical analysis of the results supports the rejection of both hypotheses and shows that AM-ART can emulate mind wandering in terms of recalling a sequence of contextually connected memory across different episodes covering a moderate subset of a person's autobiographical memory.

The rest of the paper is organized as follows. Section 2 reviews the related work. Section 3 introduces the psychological basis established by the psychologists, leading to the design of our computational AM model. Section 4 first presents the network structure of our AM model and discusses how it coincides with its psychological basis, and then introduces its dynamics with detailed formal definitions and algorithms. Section 5 introduces how our AM model emulates the wandering in reminiscence phenomenon. Section 6 presents the experimental results with visualizations and discussions. Section 7 concludes this paper and proposes future work.

2. RELATED WORK

Memory modules are always the important components of the various cognitive models. For example, cognitive models, such as Soar [13, 15], ACT-R [2], and Icarus [14], incorporate long-term memory modules besides short-term and/or working memory modules. Although the aforementioned cognitive models may not specify the exact types of the long-term memory modules they employ such as episodic [26], semantic [11], or autobiographical [21], a specific model of Soar [17] explicitly states its incorporation of an episodic memory module, which is mainly used to perform case-based reasoning by mining the stored historical data. In addition, a well-defined computational episodic memory model, which stands alone without connecting to other cognitive modules, is presented in [30]. It explicitly defines the formation, retrieval, and forgetting of the past events happened in a computer game. However, the model's usage is limited to the recall of historical data and it does not incorporate emotion [8] as one of the input fields, which is an important element in autobiographical memory [3].

Episodic memory and autobiographical memory are two closely related terms as both refer to memory collections of past events experienced by an individual. However, autobiographical memory can be considered as a special type of episodic memory containing a person's life long experience from a personal perspective [4]. Nonetheless, most existing computational episodic and autobiographical memory modules do not significantly differentiate from one another in terms of their usages and/or their representations. We identify our computational memory model as autobiographical because (a) it explicitly incorporates emotion [3] and (b) it focuses more on the cohesiveness or the connectedness among the retrieved memories. Connectedness is another criterion given in [9] to differentiate between the autobiographical and episodic memories.

Research work related to the autobiographical memory is emerging. The AM model, named Xapagy [5], is "designed to perform narrative reasoning, an activity roughly analogous to some of the mental processes humans perform with respect to stories". Xapagy is well-defined and incorporates complex natural language processing methods. However, its usage is limited to storytelling. Thudt, Baur, and Huron developed an online system to enable users to build visual mementos as a form of visualized AM based on their movement data, in the form of GPS coordinates and other context, for self-reflection and sharing experiences [25]. Pointeau, Petit, and Dominey stored the human-robot interactions as AM, by which a humanoid robot can accumulate its experience and extract regularities [19]. However, when retrieving the stored memories, all the three afore-reviewed AM models only invoke simple retrievals using the minimal amount of indexing knowledge. Specifically, they simply retrieve all memories [19], all memories of the particular user [25], or all memories comprising the selected verb [5]. Kope, Rose, and Katchabaw proposed a computational AM model using keyword-based queries for memory retrieval [12]. Taking input memories in the form of sentences describing the events happened in a game environment, their proposed model maintains a linked graph of all the parsed keywords, wherein the weights associated with the links represent the coexistence of the keyword pairs. Unlike all the computational AM models proposed in the literature, our computational AM model, named AM-ART, is designed to capture the memories comprising pictorial snapshots of one's life experience together with the associated context, namely time, location, people, activity, and emotion. It can retrieve encoded AM using different types of cues and emulate the wandering in reminiscence phenomenon of the human mind.

To the best of our knowledge, in the literature of both episodic and autobiographical memory models, there is no prior study focusing on the wandering phenomenon. Pavloski and Karimi made such an attempt to emulate the wandering effect in short- and long-term associative memories modeled by the self-trapping attractor neural network [18]. They refer wandering in their network to the mechanism of allowing "the sparsely connected network to wander to the vicinity of attractors far from the initial state." In AM-ART, we refer wandering to the process of regulated memory retrievals wherein the subsequent memories are retrieved based on highly similar but randomly mutated cues.

3. PSYCHOLOGICAL BASIS OF AM-ART

Among the various AM models established by the psychologists, the one presented by Conway and Pleydell-Pearce [7] has been accepted widely in the academic world. They categorized autobiographical memory knowledge into three levels, namely *lifetime periods, general events*, and *event-specific knowledge* (from general to specific, see Figure 1). Furthermore, they proposed that autobiographical memories can be directly accessed if the cues are specific and personally relevant. On the other hand, if the cues are general, a *generative retrieval* process must be engaged to produce more specific cues for the retrieval of relevant memories. The main difference between the direct and generative retrieval is that "the search process is modulated by control processes in generative retrieval but not, or not so extensively, in direct retrieval" [7]. This difference between the two types of memory retrieval is supported

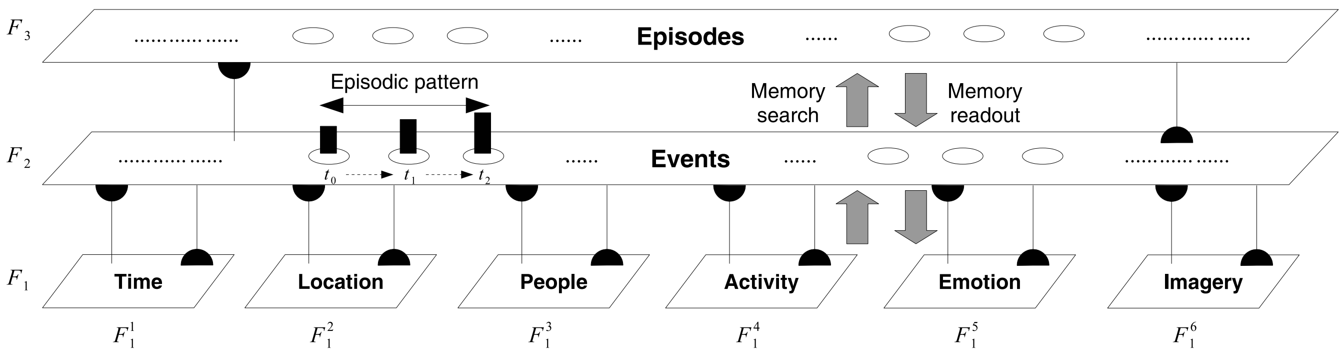

Figure 2: Network structure of AM-ART.

Table 1: Operations applied in AM-ART to realize the three stages of generative autobiographical memory retrieval.

#	Stage	Description given in [7]	AM-ART operations
1	Elaboration	"The elaboration of a cue with which to search memory and the simultaneous setting of verification criteria."	Template masking, mutation (See Algorithm 3), and setting of the vigilance parameters (see (3))
2	Strategic search	"Matching the description to records in memory."	Code activation and code competition
3	Evaluation	"Records accessed in memory were assessed against the verification criteria."	Template matching

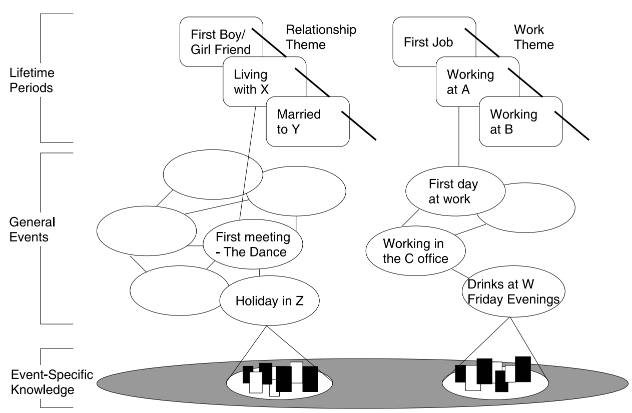

Figure 1: Illustration of the autobiographical memory hierarchy. This figure replicates Figure 1 presented in [7].

by neural imaging evidence that we can observe the difference in various regions of the brain, such as left parahippocampal gyrus, left medial prefrontal cortex, left inferior frontal gyrus, left medial parietal cortex, and left hippocampus/entorhinal cortex [1]. These established neural-psychology theories serve as the design principles of our computational autobiographical memory model.

4. AM-ART MODEL AND ITS DYNAMICS

The network structure of our computational Autobiographical Memory-Adaptive Resonance Theory (AM-ART) model is shown in Figure 2. AM-ART is a three-layer neural network that in the top-down order, its F_3, F_2, and F_1 layers encode *lifetime periods, general events*, and *event-specific knowledge*, respectively. The structure of AM-ART is consistent with the hierarchical model established by Conway and Pleydell-Pearce (see Figure 1), which serves as the theoretical basis of AM-ART. Following the examples given in Figure 1, we can highlight the correspondence between the two models. The life experience of "working at A" can be represented as a code (learned episode) in F_3 of AM-ART. The associated events of that episode, namely "first day at work", "working in the C office", and "drinks at W Friday evenings", can be represented as codes (learned events) in F_2. A specific event, taking "drinks at W Friday evenings" as an example, can be read out in F_1 that on Friday night (time), at W (location), with colleagues (people), drinking (activity), feeling happy (emotion), together with the pictorial memory (imagery).

Furthermore, memory retrieval in AM-ART replicates the three stages of the generative memory retrieval presented in [7], namely the *elaboration stage*, *strategic search stage*, and *evaluation stage*. The operations applied in AM-ART to realize the three stages of the generative memory retrieval are summarized in Table 1.

AM-ART extends the network structure of the fusion ART model [24], which is a generic self-organizing neural network comprising two layers of neural fields connected by bidirectional conditional links. However, the same bottom-up search and top-down readout operations between the layers still apply in AM-ART.

4.1 Dynamics of Fusion ART

With reference to the F_1 (comprising of six input fields) and F_2 (comprising of one association field) layers shown in Figure 2, we introduce the dynamics of fusion ART as follows.

Input vectors: Let $\mathbf{I}^k = (I_1^k, I_2^k, \ldots, I_L^k)$ denote an input vector, where I_i^k denotes the input i to channel k, for $i = 1, 2, \ldots, L$ and $k = 1, 2, \ldots, N$, where L denotes the length of channel k and N denotes the number of channels.

Input fields: Let F_1^k denote an input field that holds the input pattern for channel k. Let $\mathbf{x}^k = (x_1^k, x_2^k, \ldots, x_L^k)$ denote the activation vector of F_1^k receiving $\mathbf{I}^k$, where $x_i^k \in [0, 1]$. To prevent code proliferation when using fuzzy operations (see (1) and (4)), complement coding is applied that each $\mathbf{x}^k$ is augmented with a complement vector $\overline{\mathbf{x}}^k$ such that $\overline{x}_i^k = 1 - x_i^k$ [6]. Readers may refer to [28] for a more detailed description of complement coding.

Association field: Let $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_C\}$ denote the activation vector of F_2, where C denotes the number of codes in F_2.

Weight vectors: Let $\mathbf{w}_j^k$ denote the weight vector of the jth code in F_2 for learning the input pattern in F_1^k.

Parameters: The dynamics of fusion ART are regulated by the parameters associated with all the input fields in the lower layer, namely choice parameters $\alpha^k > 0$, learning rate parameters $\beta^k \in [0, 1]$, contribution parameters $\gamma^k \in [0, 1]$, where $\sum \gamma^k = 1$, and vigilance parameters $\rho^k \in [0, 1]$.

Code activation: A bottom-up knowledge search starts from the computation of the activations (choice function values) for all codes in F_2. Specifically, given $\mathbf{x}^k$, for each F_2 code j, the choice function T_j is computed as follows:

$$T_j = \sum_k \gamma^k \frac{|\mathbf{x}^k \wedge \mathbf{w}_j^k|}{\alpha^k + |\mathbf{w}_j^k|}, \tag{1}$$

where the fuzzy AND operation $\wedge$ is defined by $p_i \wedge q_i \equiv min(p_i, q_i)$ and the norm $|.|$ is defined by $|\mathbf{p}| \equiv \sum_i p_i$.

Code competition: A code competition process follows under which the F_2 code with the highest choice function value is identified. The winner is indexed at J, where

$$T_J = \arg\max\{T_j : \text{for all } F_2 \text{ code } j\}. \tag{2}$$

Template matching: A template matching process checks whether resonance occurs at the chosen code J. Specifically, resonance occurs if for each channel k, the match function m_J^k meets its vigilance criterion, such that

$$m_J^k = \frac{|\mathbf{x}^k \wedge \mathbf{w}_J^k|}{|\mathbf{x}^k|} \geq \rho^k. \tag{3}$$

If any vigilance constraint is violated, mismatch reset occurs in which T_J is set to 0 for the duration of the input presentation. Thus, another F_2 code will be selected as the new winner. This search and evaluation process is guaranteed to end, because either a committed code that satisfies the vigilance criteria will be identified or an uncommitted one (all weight values initialized to 1, which definitely satisfies the criteria) will be recruited to encode the new input pattern. Once an uncommitted code is recruited, a new uncommitted code will be autonomously added in F_2. Thus, the fusion ART model self-organizes its network structure.

Template learning: Once the code J is identified as wherein resonance occurs, for each channel k, the weight vector $\mathbf{w}_J^k$ can be updated by the following learning rule:

$$\mathbf{w}_J^{k(\text{new})} = (1 - \beta^k)\mathbf{w}_J^{k(\text{old})} + \beta^k(\mathbf{x}^k \wedge \mathbf{w}_J^{k(\text{old})}). \tag{4}$$

Knowledge readout: This top-down retrieval procedure is invoked when the chosen F_2 code J presents its weight vectors to the input fields in F_1, such that $\mathbf{x}^{k(\text{new})} = \mathbf{w}_J^k$.

Template masking: Due to the dynamics of fusion ART, not all input vectors have to be presented for knowledge retrieval [23]. In such cases, all the values of the absent vector $\mathbf{x}^k$ (including complements) are set to 1.

4.2 Encoding and Retrieval of Events

In AM-ART (see Figure 2), the input fields in F_1 encode the 5W1H, respectively. To make the activation vectors $\mathbf{x}^k$ compact and generic, we use normalized values to represent *time* and *location*. On the other hand, we use categorical values to present *people*, *activity*, *emotion*, and *imagery*.

Time vector $(\mathbf{x}^1)$**:** It represents *when* the event happened in terms of six values, namely the normalized day ($x_1^1 = I_1^1/31$), month ($x_2^1 = I_2^1/12$), and year ($x_3^1 = (I_3^1 - 1900)/200$), together with their complements.

Location vector $(\mathbf{x}^2)$**:** It represents *where* the event happened in terms of four values, namely the normalized latitude ($x_1^2 = (I_1^2 + 90)/180)$) and longitude ($x_2^2 = (I_2^2 + 180)/360$) ($\mathbf{I}^2$ denotes the input vector of location, which is determined via Google Geocoder API), together with their complements.

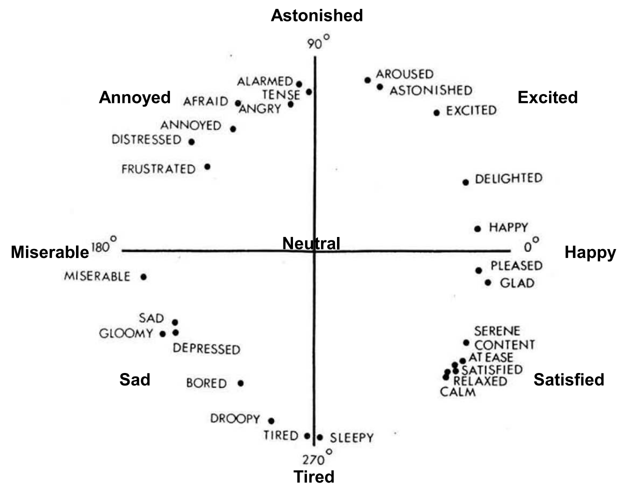

Figure 3: Categorization of emotion based on Russell's circumplex model of affect [22].

People vector $(\mathbf{x}^3)$**:** It is a binary-valued vector representing *who* were involved in the event. Its length corresponds to the categorization of people based on relationship.

Activity vector $(\mathbf{x}^4)$**:** It is a binary-valued vector representing *what* was the event. Similarly, its length corresponds to the categorization of activities.

Emotion vector $(\mathbf{x}^5)$**:** It is a binary-valued vector representing *how* was the feeling during the event. Emotion is an important component of our past experience, which affects the encoding and retrieval of autobiographical memories [3]. We categorize nine types of emotion, namely neutral, astonished, excited, happy, satisfied, tired, sad, miserable, and annoyed (see Figure 3). Thus, the length of x^5 is 18. This categorization follows the pleasure-arousal model established by Russell [22] and applied in [27].

Imagery vector $(\mathbf{x}^6)$**:** It is a binary-valued vector representing *which* pictorial memory was associated with the event. Its value encodes the file-path of the stored imagery. During memory retrieval, this vector is not presented along with the others as a part of the retrieval cue. Therefore, in Algorithms 3 and 4, only the first five input fields of F_1 (see Figure 2) are involved.

The F_2 layer of AM-ART encodes events. The procedure of event encoding and retrieval is shown in Algorithm 1.

Algorithm 1 Event encoding and retrieval

```
Encode x^k in F_1 w.r.t the given input pattern I^k
Activate all codes in F_2                              ▷ see (1)
repeat selecting the winner code J                     ▷ see (2)
until resonance occurs                                 ▷ see (3)
if encoding is required then perform learning          ▷ see (4)
end if
if retrieval is required then read out w_J^k in F_1
end if
```

4.3 Encoding and Retrieval of Episodes

Assume the related events of one episode happened at $t_0, t_1, \ldots, t_n$ and let y_{t_i} denote the activation value of the event happened at t_i. To encode the sequence of the events, we need to always hold the inequality that $y_{t_n} > y_{t_{n-1}} > \cdots > y_{t_0}$. Therefore, we use a decay parameter $\tau \in (0, 1)$ to regulate the activation decays, such that $y_j^{(\mathrm{new})} = y_j^{(\mathrm{old})}(1 - \tau)$ at each new time step.

The F_3 layer of AM-ART encodes episodes to associate the related events encoded in F_2. The procedure of episode encoding and retrieval is shown in Algorithm 2.

Algorithm 2 Episode encoding and retrieval

for all subsequent events of an episode **do**
  select the winner code J in F_2 w.r.t $\mathbf{x}^k$ in F_1
  $\mathbf{y}_J \leftarrow 1$ ▷ or a predefined value if using partial sequence to identify the episode
  for all previously selected codes in F_2 **do**
    $\mathbf{y}_i^{(\mathrm{new})} \leftarrow \mathbf{y}_i^{(\mathrm{old})}(1 - \tau)$
  end for
end for
Select the winner code J' in F_3 w.r.t $\mathbf{y}$
if encoding is required **then** learn the weight vector $\mathbf{w}'_{J'}$ in F_3: $\mathbf{w}'^{(\mathrm{new})}_{J'} \leftarrow (1 - \beta_2)\mathbf{w}'^{(\mathrm{old})}_{J'} + \beta_2(\mathbf{y} \wedge \mathbf{w}'^{(\mathrm{old})}_{J'})$
end if
if retrieval is required **then** read out $\mathbf{w}'_{J'}$ in F_2
end if

5. WANDER IN REMINISCENCE

In this section, we present how we enable AM-ART to wander in reminiscence. The wandering involves two major procedures, namely mutating the retrieval cue and iteratively retrieving autobiographical memories using a mutated cue in each iteration.

We show the procedure of mutating a retrieval cue in Algorithm 3, which is conceptually similar to mutate the chromosomes in genetic algorithms [10]. We can summarize the mutation procedure as that upon given a retrieval cue, we intentionally add regulated noise to it in the randomly determined positions. The mutation process is regulated by two parameters, namely mutation rate $T \in [0, 1]$ and noise level $L \in [0, 1]$. We use mutation rate T to control the probability on mutating the value of the retrieval cue in a certain field and use noise level L to control the amount of mutation.

Algorithm 3 Mutation of Memory Retrieval Cues

Given a memory retrieval cue: $\mathbf{x} = \{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^5\}$
for all $\mathbf{x}^i \in \mathbf{x}$ **do**
  if $rand() \leq T$ **then** ▷ the ith field is selected for mutation, where $0 \leq rand() < 1$
    for all x_j^i, where $j = 1, 2, \ldots, \frac{|\mathbf{x}^i|}{2}$ **do** ▷ $|\mathbf{p}|$ denotes the cardinality of vector $\mathbf{p}$
      if $i \leq 2$ && $rand() \leq \frac{2}{|\mathbf{x}^i|}$ **then** ▷ for normalized vector, the jth value is selected for mutation
        $x_j^i \leftarrow (1 + (rand() - 0.5)L) \cdot x_j^i$;
        bound x_j^i within the $[0, 1]$ interval
      else if $i \geq 3$ && $rand() \leq \frac{2(1+L)}{|\mathbf{x}^i|}$ **then** ▷ for binary-valued vector
        negate the binary value of x_j^i
      end if
    end for
    for all x_j^i, where $j = \frac{|\mathbf{x}^i|}{2} + 1, \frac{|\mathbf{x}^i|}{2} + 2, \ldots, |\mathbf{x}^i|$ **do** $x_j^i \leftarrow 1 - x^i_{j-|\mathbf{x}^i|/2}$ ▷ recompute the complements
    end for
  end if
end for

Table 2: List of AM-ART parameters used in experiments.

Parameter	Value	Description/Remark
Choice (α^k)	0.1	Mainly used to avoid having NaN in (1)
Learning rate (β^k)	1	Not in use during memory retrieval
Contribution (γ^k)	0.167	Equally assigned, such that $\sum \gamma^k = 1$
Vigilance (ρ^k)	1	Value varies in different retrieval settings, see Fig. 4; set to 0 during wandering, see Algo. 4
Decay rate (τ)	0.1	Used for encoding event sequence

We show how AM-ART wanders in reminiscence in Algorithm 4. We can summarize the wandering procedure as that upon given the memory retrieval cue at the beginning of each iteration, we first use the given cue to retrieve the most related event, which has not been included in the retrieved memory set, and then we append the retrieved event to the memory set and proceed to the next iteration with the mutated retrieval cue. The stopping criterion of Algorithm 4 is by retrieving a predefined number (N) of events, which is mainly set for experimental purposes. However, this criterion can be easily removed or modified for continuous retrieval.

Algorithm 4 Wandering in Reminiscence

Given a memory retrieval cue: $\mathbf{x} = \{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^5\}$
$\rho^k \leftarrow 0$, for $k = 1, 2, \ldots, 6$ ▷ remove all vigilance criteria during memory retrieval
$\mathbf{M} = \emptyset$; ▷ initialize the memory set
repeat
  repeat identify E in F_2 w.r.t $\mathbf{x}$; ▷ find winner event
    $\mathbf{y}_E \leftarrow 0$ ▷ suppress its activation value
  until $E \notin \mathbf{M}$
  $\mathbf{M} \leftarrow \mathbf{M} \cup \{E\}$ ▷ retrieving order preserved in $\mathbf{M}$
  Mutate $\mathbf{x}$ ▷ see Algorithm 3
until $|\mathbf{M}| = N$
retrieve all events in $\mathbf{M}$

6. EXPERIMENTS

To conduct experiments, we collected a data set comprising 53 snapshots of events, together with the corresponding context, of Mr. Obama, the current President of the USA. The 53 events are organized in twelve episodes, each of which contains three to seven events. Most of these photos are collected from Zimbio[1] and others via Google Images. From the online web pages, we directly extracted all features except emotion, which was manually derived from the picture and its context. It is important to note that the automated extraction of such contextual information is not the focus of this paper, but can be considered as a future extension to AM-ART.

According to the collected data set, we define eight types of relationship among people, namely family, neighbors, spouse, friends, classmates, colleagues, acquaintances, and strangers. Moreover, we define fifteen classes of activities, namely meal, leisure, travel, holiday, shopping, night-out, recreation, sports, exercise, work, gathering, party, celebration, wedding, and school. After formalizing the input vectors, we present the data samples to AM-ART to encode the 53 events in F_2 and the 12 episodes in F_3 (see Algorithms 1 and 2, respectively).

The parameters of AM-ART used in all experiments are listed in Table 2. Most such parameters take on a standard set of parameter values and all do not require tuning in the experiments.

[1]URL: http://www.zimbio.com/Barack+Obama/pictures/pro

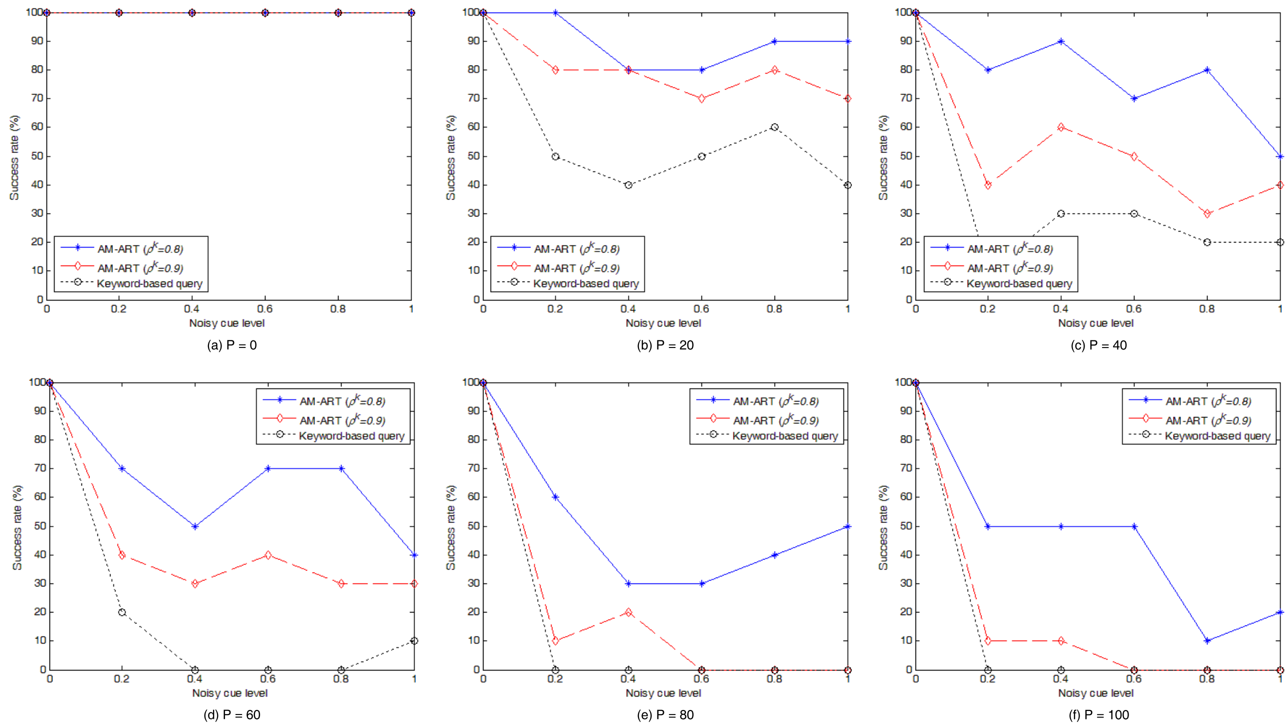

Figure 4: Success rates of memory retrievals by AM-ART and the keyword-based query method in response to noisy cues with different cue completeness percentile P and different noisy cue level L'.

6.1 Autobiographical Memory Retrieval Using Exact, Partial, and Noisy Cues

After AM-ART encodes the memories, we first test its performance in memory retrieval using the following three types of cues:

Exact cues: We randomly select an event from the data set and use its representation vector $\mathbf{x} = \{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^5\}$ (except the imagery path) as the *exact* retrieval cue.

Partial cues: Given an exact cue $\mathbf{x}$ and a cue completeness percentile P, in the *partial* cue $\mathbf{x}_P$, for each field k, where $k = 1, 2, \ldots, 5$, if $rand() \leq P/100$, $\mathbf{x}_P^k = \mathbf{x}^k$. Otherwise, $\mathbf{x}_P^k = \{1, \ldots, 1\}$. In ART models, a vector consisting of all 1s (including the complements) signifies do-not-care. When $P = 100$, $\mathbf{x}_P = \mathbf{x}$.

Noisy cues: Given a partial cue $\mathbf{x}_P$ and a noisy cue level L', in the selected fields of $\mathbf{x}_P$, where $\mathbf{x}_P^k \neq \{1, \ldots, 1\}$, we introduce noise to obtain the *noisy* cue $\mathbf{x}_N$. The process of introducing noise to a retrieval cue exactly follows the process from Step 4 to Step 13 described in Algorithm 3, wherein $L = L'$.

For the benchmarking method, we select the keyword-based query method, which is employed by many existing photo or memory repositories and the recently proposed computational AM models [12]. Its retrieval criterion is based on whether the given retrieval cues exactly match the corresponding portions of the stored records. In the context of this paper, a set of events will be retrieved after performing the keyword-based query in response to the given retrieval cue. If the event initially being used to generate the given cue can be found in the retrieved set, we consider the retrieval as a success. Otherwise, we consider it as a failure.

For AM-ART to handle the noisy cues, we can lower the vigilance parameters ρ^k. Actually, in the front-end of our online AM agent [29], we allow users to define the confidence level (which refers to ρ^k in the back-end) associated with any field of the retrieval cue they provide. In such a way, the users can retrieve certain memories even by providing imperfect cues. This feature is the major advantage of our agent over many other existing photo or memory repository type of applications. In response to the given retrieval cue, our AM-ART agent retrieves a set of memories and plays them back to the users. For fair comparisons, the judging criterion between a successful retrieval and a failure is the same as that of the keyword-based query method.

For each experiment, we repeat for ten times, wherein the retrieval cue is randomly selected or generated at each time, and we aggregate the results for further analysis.

Due to the way we use to generate partial cues and the criterion we adopt to count successful retrievals, both AM-ART and the keyword-based query method achieve 100% success rate in response to exact and partial retrieval cues. However, dealing with noisy cues is challenging because uncertainties can be well handled by our human brains but not by many computational models. The experimental results on memory retrievals in response to noisy cues are visualized in Figure 4. Please note that when $\rho^k = 1$, which means AM-ART requires exact matches, the performance of AM-ART is equivalent to that of the keyword-based query method.

The six sub-figures of Figure 4 show the success rates according to different values of the cue completeness percentile P. When $P = 0$, all models essentially retrieve all the events in the data set. Therefore, in Figure 4(a), all models achieve 100% success rate. As the value of P increases, the retrieval cue consists of an increasing number of input fields. Therefore, as expected, we observe that the performance of all models generally decline as the values of the cue completeness percentile P and the noisy cue level L' increase.

By comparing the performance of the AM-ART models against that of the keyword-based query as shown in Figure 4, AM-ART

Table 4: **A partial subsequence of the autobiographical memories retrieved by AM-ART (W22) during wandering.**

#	Time (when)	Location (where)	People (who)	Activity (what)	Emotion (how)	Event # of Episode #
Cue	**08-Jul-2014**	Honolulu, United States	**friends**	**night-out**	**happy**	randomly initialized
1	**08-Jul-2014**	Wynkoop Brewing Co.	**friends & acquaintances**	**night-out**	**happy**	Event #1 of Episode #5
2	**10-Dec-2014**	**Joint Base Anacostia-Bolling**	**friends & acquaintances & strangers**	**party**	**happy**	Event #3 of **Episode #10**
3	**10-Dec-2014**	**Joint Base Anacostia-Bolling**	**friends & acquaintances & strangers**	**party**	astonished	Event #1 of **Episode #10**
4	08-Jul-2014	Wynkoop Brewing Co.	**friends & acquaintances**	night-out	happy	Event #2 of Episode #5

*Associated values across consecutive retrievals are highlighted in **bold**. **Retrieved events #1 and #4 differ in their event imagery.

(1) Event #1 of Episode #5
(2) Event #3 of Episode #10
(3) Event #1 of Episode #10
(4) Event #2 of Episode #5

Figure 5: The imagery playback of the retrieved events following the order presented in Table 4.

Table 3: Naming conventions on the parameter settings.

Name	Setting (T, L)	Name	Setting (T, L)
S	Sequential retrieval	R	Random retrieval
WSS	Wandering (0.0, 0.0)	WRR	Wandering (1.0, 1.0)
W11	Wandering (0.2, 0.1)	W12	Wandering (0.2, 0.2)
W21	Wandering (0.4, 0.1)	W22	Wandering (0.4, 0.2)

clearly performs better in terms of the successful retrieval rate in response to noisy cues. This finding is encouraging because it suggests that by lowering its vigilance when handling the noisy cues, AM-ART can better deal with the uncertainty involved in the retrieval process, which is more towards the pursuit of human-like intelligence in terms of handling imperfect information.

6.2 Wandering in Reminiscence

As listed in Table 3, different value combinations of the mutation rate T and the noise level L (see Algorithm 3) are used in different settings for performance comparisons on wandering in reminiscence. We choose $T = 0.2$ or 0.4 such that during mutation, we can expect one or two fields of the cue to be mutated on average, respectively. We choose $L = 0.1$ or 0.2 to generate a reasonably small amount of noise during wandering. We also test the extreme cases (WSS and WRR), wherein T and L are set to the boundary values. The experimental results obtained using these parameter settings are compared against each other and against the two baselines, namely sequential retrieval by strictly retrieving memories according to the event encoding order during learning and random retrieval, wherein all (unique) memories are retrieved randomly.

In Table 4, we present a partial sequence of the memories (see Figure 5) retrieved by AM-ART during wandering in the W22 configuration. Table 4 well demonstrates that AM-ART wanders off Episode #5 (consisting of five events) to Episode #10 (consisting of seven events), but wanders back to Episode #5 after two steps.

We conducted series of experiments and in each run we retrieved N events (see Algorithm 4). In this paper, we set $N = 35$, which is around 2/3 of the total number of events (53) in the data set. To numerically evaluate the performance of AM-ART during wandering, we devise the following three measurement metrics, which are all normalized within the range of [0,1]. A higher value of each metric only indicates a better performance in the respective aspect.

Coverage C**:** It measures the distribution of the retrieved events among all the episodes. Let S_C denote the number of episodes, which have at least one event in the retrieved memory set $\mathbf{M}$, and $S_{C_{\min}}$ denote the minimum value of S_C, which can be determined by N and the composition of the data set in use. Thus, $C = (S_C - S_{C_{\min}})/(P - S_{C_{\min}})$, where P denotes the total number of episodes in the data set. For this data set, $P = 12$ and $S_{C_{\min}} = 8$.

Diversity D**:** It measures the diversity of the retrieved memories. Let $\mathbf{u} = \{u_1, u_2, \ldots, u_{N-1}\}$ denote a binary-valued vector to capture the relationship among the subsequently retrieved events in $\mathbf{M}$. For $k = 1, 2, \ldots, N-1$, if $\mathbf{M}_k$ and $\mathbf{M}_{k+1}$ are retrieved from different episodes, $u_k = 1$. Otherwise, $u_k = 0$. Similar to $S_{C_{\min}}$, $S_{D_{\min}}$ denoting the minimum number of 1s in $\mathbf{u}$ can be determined by N and the composition of the data set. Thus, $D = (H(\mathbf{u}) - S_{D_{\min}})/(N - 1 - S_{D_{\min}})$, where $H(\mathbf{u})$ computes the Hamming weight of $\mathbf{u}$. For this data set, $S_{D_{\min}} = 7$.

Relatedness R**:** It measures the relatedness of the subsequently retrieved memories. Let $\mathbf{v}_j = \{\mathbf{v}_j^1, \mathbf{v}_j^2, \ldots, \mathbf{v}_j^5\}$ denote the jth retrieved event in vector form (without complements), where $j = 1, 2, \ldots, N$. Moreover, let E_k, for $k = 1, 2, \ldots, N-1$, denote the Euclidean distance (dissimilarity) between $\mathbf{v}_j$ and $\mathbf{v}_{j+1}$. Thus, $R = (S_{R_{\max}} - \sum E_k)/S_{R_{\max}}$, where $S_{R_{\max}}$ denotes the maximum dissimilarity among all the retrieved memories. For this data set, we estimate $S_{R_{\max}}$ as follows. The length (without compliments) of the first five input fields (see Figure 2) is 2, 3, 8, 15, and 9, respectively. However, for activity and emotion vectors, one and only one bit is 1. Therefore, the maximum dissimilarity between any two consequent retrievals is $\sqrt{3 \cdot \sqrt{2} + \sqrt{3} + \sqrt{8}} \approx 8.8$. Thus, $S_{R_{\max}} = 34 \times 8.8 = 299.2$.

For fair comparisons, all experimental settings use the same randomly initialized cue to retrieve the first piece of autobiographical memory and each experiment is repeated for ten times for statistical analysis. To test whether AM-ART can simulate wandering in reminiscence, we propose the following two null hypotheses.

$\mathbf{H_{01}}$**:** The performance of AM-ART during wandering is at the same level as that of sequential memory retrieval.

$\mathbf{H_{02}}$**:** The performance of AM-ART during wandering is at the same level as that of random memory retrieval.

We visualize the averaged experimental results in Figure 6. Across the three sub-figures, it seems that the performance of typical wandering settings, namely W11, W12, W21, and W22, is always in the middle range and appears to be different from the rest. However, we need to perform statistical analysis for further investigations.

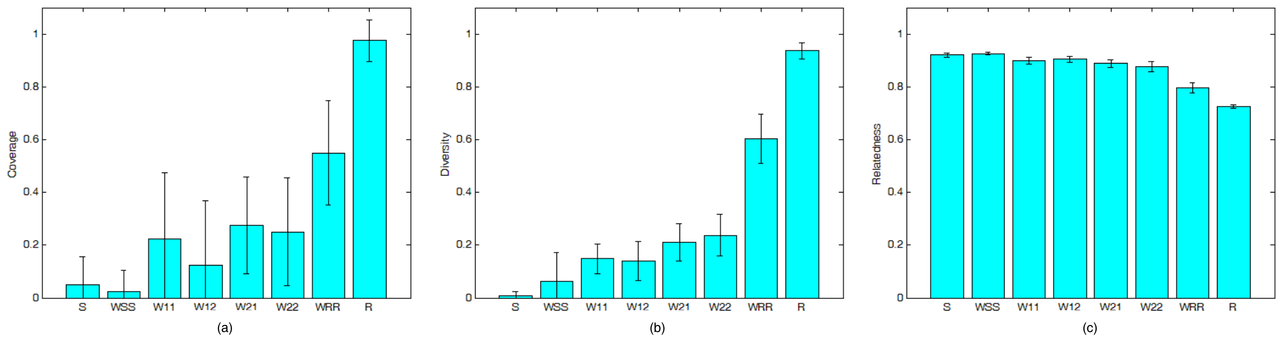

Figure 6: Performance comparisons of various configurations in terms of (a) Coverage C; (b) Diversity D; and (c) Relatedness R.

Table 5: Statistical analysis using two-tail p-values. If $p < 0.05$, the value is highlighted in bold.

Model	S			WSS			WRR			R		
	Cover.	**Diver.**	**Relat.**	**Cover.**	**Diver.**	**Relat.**	**Cover.**	**Diver.**	**Relat.**	**Cover.**	**Surpr.**	**Relat.**
W11	0.110	**<0.001**	**<0.001**	0.053	0.086	**<0.001**	**0.006**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**
W12	0.279	**<0.001**	**<0.001**	0.104	0.110	**0.001**	**0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**
W21	**0.004**	**<0.001**	**<0.001**	**0.001**	**0.008**	**<0.001**	**0.003**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**
W22	**0.022**	**<0.001**	**<0.001**	**0.010**	**0.007**	**<0.001**	**0.006**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**

Table 6: Performance evaluation of the typical settings.

	W11	**W12**	**W21**	**W22**
A-mean of C, D, and R	0.424	0.390	**0.458**	0.454
G-mean of C, D, and R	0.311	0.252	**0.375**	0.373

*The highest value in each row is highlighted in **bold**.

Based on the statistical analysis results shown in Table 5, we observe that (a) all typical wandering settings behave significantly different from that of random memory retrievals (WRR and R) and (b) with mutation rate $T = 0.4$, the typical wandering settings behave significantly different from those of both sequential (S and WSS) and random memory retrievals. The latter finding is further supported in Table 6, wherein in terms of both arithmetic and geometric means, W21 and W22 perform better than W11 and W12.

In summary, with a mutation rate of $T = 0.4$, we can reject both $\mathbf{H_{01}}$ and $\mathbf{H_{02}}$ with a strong statistical support. Based on the various performance comparisons, we find that W21 configuration performs best among all the typical settings in terms of the overall measure (see Table 6) and it is truly different from both sequential and random retrievals (see Table 5). Therefore, we show that AM-ART in the W21 configuration effectively retrieves a moderate subset of an individual's autobiographical memory in a contextually connected sequence across multiple episodes and can simulate the phenomenon of wandering in reminiscence.

7. CONCLUSION

This paper has presented a computational model, named AM-ART, for the encoding and retrieval of autobiographical memories in online autonomous agents modeling users' life experiences. To the best of our knowledge, AM-ART is the first computational model simulating the mind wandering phenomenon in human autobiographical memories. The network structure and the dynamics of AM-ART follow the AM model established by psychologists [7], which has been supported by neural imaging evidence [1]. Specifically, the three-layer AM-ART structure coincides with the hierarchy of the AM model presented in [7] and the operations of AM-ART coincide with the retrieval mechanisms presented in [7]. In terms of memory retrieval in response to noisy cues, AM-ART is shown as performing better than the keyword-based query method, which cannot handle noisy cues in many existing photo or memory repositories. To evaluate the performance of AM-ART in wandering in reminiscence, we devise three measurement metrics, namely coverage, diversity, and relatedness. Statistical analysis of the experimental results shows that AM-ART can emulate mind wandering in terms of recalling a sequence of contextually connected memory across different episodes covering a moderate subset of a person's autobiographical memory.

One important feature of AM not discussed in this paper is that humans partially reconstruct their memories based on their internal mental states, such as self-intention, central concerns, or personal characteristics [4]. Therefore, one key direction of our future work is to introduce such factors into AM-ART to simulate the reconstruction of AM from the personal perspective. One possible approach is to experiment with different parameter values, including the mutation rate and the noise level, associated with each field representing the 5W1H according to the user's biases and preferences.

We show in this paper that AM-ART is designed in accordance to a well-established psychological basis and it can encode and retrieve real-life autobiographical memory, comprising pictorial snapshots of one's life experiences together with the associated context. However, the performance of AM-ART in wandering in reminiscence is only shown through a statistical analysis of the experimental results. Going forward, we shall conduct user studies on the human-likeness of the memory retrievals and whether the overall system can deliver enjoyment and desirably health improvement through memory playback [20].

Acknowledgments

We thank Wei-Xuan Huang for her contribution in data collection and system integration. This research is supported in part by the National Research Foundation Singapore under its Interactive Digital Media (IDM) Strategic Research Programme.

REFERENCES

[1] D. R. Addis, K. Knapp, R. P. Roberts, and D. L. Schacter. Routes to the past: Neural substrates of direct and generative autobiographical memory retrieval. *Neuroimage*, 59(3):2908–2922, 2012.

[2] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin. An integrated theory of the mind. *Psychological Review*, 111:1036–1060, 2004.

[3] D. Berntsen and D. C. Rubin. Emotionally charged autobiographical memories across the life span: The recall of happy, sad, traumatic, and involuntary memories. *Psychology and Aging*, 17(4):636–652, 2002.

[4] S. Bluck and L. J. Levine. Reminiscence as autobiographical memory: A catalyst for reminiscence theory development. *Ageing and Society*, 18(2):185–208, 1998.

[5] L. Boloni. Integrating perception, narrative, premonition and confabulatory continuation. *Biologically Inspired Cognitive Architectures*, 8:120–131, 2014.

[6] 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*, 4(6):759–771, 1991.

[7] M. A. Conway and C. W. Pleydell-Pearce. The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107(2):261–288, 2000.

[8] J. Dias and A. Paiva. Feeling and reasoning: A computational model for emotional characters. *Progress in artificial intelligence, Lecture Notes in Computer Science*, 3808:127–140, 2005.

[9] A. Gilboa. Autobiographical and episodic memory–one and the same? Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia*, 42(10):1336–1349, 2004.

[10] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.

[11] J. Hart and M. A. Kraut, editors. *Neural Basis of Semantic Memory*. Cambridge University Press, 2007.

[12] A. Kope, C. Rose, and M. Katchabaw. Modeling autobiographical memory for believable agents. In *Proceedings of AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, pages 23–29, 2013.

[13] J. E. Laird. *The Soar Cognitive Architecture*. MIT Press, 2012.

[14] P. Langley. Cognitive architectures and general intelligent systems. *AI Magazine*, 27:33–44, 2006.

[15] J. Li and J. Laird. Spontaneous retrieval from long-term memory for a cognitive architecture. In *Proceedings of AAAI Conference on Artificial Intelligence*, pages 544–550, 2015.

[16] Y.-C. Lin, Y.-T. Dai, and S.-L. Hwang. The effect of reminiscence on the elderly population: A systematic review. *Public Health Nursing*, 20(4):297–306, 2003.

[17] A. M. Nuxoll and J. E. Laird. Extending cognitive architecture with episodic memory. In *Proceedings of AAAI Conference on Artificial Intelligence*, pages 1560–1565, 2007.

[18] R. Pavloski and M. Karimi. Wandering mind: A self-trapping network can converge to attractors far from the initial state. In *Proceedings of International Joint Conference on Neural Networks*, pages 1168–1173, 2001.

[19] G. Pointeau, M. Petit, and P. F. Dominey. Successive developmental levels of autobiographical memory for learning through social interaction. *IEEE Transactions on Autonomous Mental Development*, 6(3):200–212, 2014.

[20] C. Routledge, T. Wildschut, C. Sedikides, and J. Juhl. Nostalgia as a resource for psychological health and well-being. *Social and Personality Psychology Compass*, 7(11):808–818, 2013.

[21] D. C. Rubin, editor. *Autobiographical Memory*. Cambridge University Press, 1988.

[22] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[23] A.-H. Tan. Direct code access in self-organizing neural networks for reinforcement learning. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 1071–1076, 2007.

[24] A.-H. Tan, G. A. Carpenter, and S. Grossberg. Intelligence through interaction: Towards a unified theory for learning. In *Proceedings of International Symposium on Neural Networks*, pages 1094–1103, 2007.

[25] A. Thudt, D. Baur, and S. Huron. Visual mementos: Reflecting memories with personal data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):369–378, 2016.

[26] E. Tulving. *Elements of Episodic Memory*. Oxford University Press, 1983.

[27] D. Wang and A.-H. Tan. Mobile humanoid agent with mood awareness for elderly care. In *Proceedings of International Joint Conference on Neural Networks*, pages 1549–1556, 2014.

[28] D. Wang and A.-H. Tan. Creating autonomous adaptive agent in a real-time first-person shooter computer game. *IEEE Transactions on Computational Intelligence and AI in Games*, 7(2):123–138, 2015.

[29] D. Wang and A.-H. Tan. MyLife: An online personal memory album. In *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 243–244, 2015.

[30] 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*, 23(10):1574–1586, 2012.