Partial Forgetting in Episodic Memory for Virtual Human

Juzheng Zhang
Supervisors: Jianmin Zheng, Nadia Magnenat Thalmann
School of Computer Engineering

23/01/2014
Definition for EM

- Episodic Memory [1]
  - Episodes or Events
    - Episode $\rightarrow$ Sequence of events with temporal-spatial relations
    - Data Structure $\rightarrow$ (Time, Context, Content)
  - Encode $\rightarrow$ Store (Forgetting) $\rightarrow$ Retrieve
  - Short-term Memory $\rightarrow$ Long-term memory

Forgetting vs. Partial Forgetting

- **Forgetting**
  - Entire episode
  - Total remembered or forgotten
  - Most previous works

- **Partial Forgetting**
  - Features in episode
  - Some are remembered, others are forgotten
  - No work [1]

States-of-Arts

• **Forgetting of EM in AI (problem-solving)**
  – Activation factors (at encoding)
    • Utility
  – Updating factors (after encoding)
    • Frequency of use (i.e. Times of being retrieved)
    • Temporal decay

• **Forgetting of EM in VC (communication)**
  – Activation factors
    • Utility or Emotional salience
    • Goal-related event
  – Updating factors (The same)
States-of-Arts

• **Problems**
  1. Subjective (*need predefinitions*)
  2. Task-dependent (e.g. goal-related event)
  3. No partial forgetting (*assign utility for every feature?*)
  4. Not suitable for EM in VC (*communication ↔ information*)
Objective

• Propose Partial Forgetting Algorithm
  – Objective (no predefinition)
  – General (task-independent)
  – Support partial forgetting (build activations for features)
  – Suitable for EM in VC (focus on how much information each feature carries)
Frame Work of Partial Forgetting
Compute Intensity

- **Theory**
  - If a context or content of an episode is
    1. frequently happened
    2. similar to contexts or contents stored in memory
  Then this context or content is easy to be remembered (more activated), and vice versa.
Compute Intensity $I_i$

• Theory

- If a feature is frequently happened, compared with other features in the same context or content of the episode, this feature is:
  1. easy to be remembered (more activated),
  2. has less contribution to the retrieval of the episode (more competition),

and vice versa.
Update Intensity

- **Theory**
  - Time decay: Ebbinghaus forgetting curve [1]
  - Reinforcement: retrieve

\[
I^t = I^0 \cdot (1 - e^{-\frac{t}{\tau}})
\]

Experiment

• New comer

• Initialization
  – The personal information of 20 members

• Encoding
  – Name (e.g. “Zhang”)
  – Nationality (e.g. “Chinese”)
  – School (e.g. “SCE”)
  – Position (e.g. “Phd”)

• Retrieval
  – Retrieval cue (e.g. “Zhang, #, #, Phd”)
Experiment
## Experiment

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Nationality</th>
<th>School</th>
<th>Position</th>
<th>Meeting times</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nadia</td>
<td>Swiss</td>
<td>IMI</td>
<td>Director</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Daniel</td>
<td>Swiss</td>
<td>IMI</td>
<td>Professor</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Andy</td>
<td>Singapore</td>
<td>EEE</td>
<td>Professor</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Cai</td>
<td>Singapore</td>
<td>MAE</td>
<td>Professor</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Cai</td>
<td>Chinese</td>
<td>SCE</td>
<td>Professor</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>He</td>
<td>Chinese</td>
<td>SCE</td>
<td>Professor</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Yuan</td>
<td>Chinese</td>
<td>EEE</td>
<td>Professor</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Zheng</td>
<td>Chinese</td>
<td>SCE</td>
<td>Professor</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Xiao</td>
<td>Chinese</td>
<td>IMI</td>
<td>Fellow</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Zerrin</td>
<td>Turkey</td>
<td>IMI</td>
<td>Fellow</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Zhang</td>
<td>Chinese</td>
<td>IMI</td>
<td>Fellow</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Ren</td>
<td>Singapore</td>
<td>EEE</td>
<td>Fellow</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>Shantanu</td>
<td>Indian</td>
<td>IMI</td>
<td>Fellow</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>Li</td>
<td>Chinese</td>
<td>MAE</td>
<td>Phd</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>Wu</td>
<td>Chinese</td>
<td>EEE</td>
<td>Phd</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>Zhang</td>
<td>Chinese</td>
<td>SCE</td>
<td>Phd</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>Hou</td>
<td>Chinese</td>
<td>EEE</td>
<td>Phd</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>Cai</td>
<td>Chinese</td>
<td>MAE</td>
<td>Phd</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>Dong</td>
<td>Chinese</td>
<td>SCE</td>
<td>Phd</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>Yuan</td>
<td>Chinese</td>
<td>SCE</td>
<td>Phd</td>
<td>1</td>
</tr>
<tr>
<td>new</td>
<td>Wang</td>
<td>Chinese</td>
<td>SPMS</td>
<td>Professor</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix: Details
EM in AI vs. EM in VC

• EM in Artificial Intelligence [1]
  – Objective
    • Improve the performance of the agent to accomplish certain tasks.
  – Focus
    • Goal/Task
  – Method
    • Similar to Case-based Reasoning (CBR)

EM in AI vs. EM in VC

• EM in Virtual Companion [1]
  – Objective
    • Communicate with users with aware of past experience
  – Focus
    • Information
  – Method
    • Information management & retrieval

[1] What does your actor remember? towards characters with a full episodic memory
# General Data Structure of Episodes

<table>
<thead>
<tr>
<th>Time</th>
<th>Physical time (e.g. 17:17:11 13-02-2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel Features (Context)</td>
<td>Environment Features</td>
</tr>
<tr>
<td></td>
<td>✓ Outside Features</td>
</tr>
<tr>
<td></td>
<td>• (Narrative time, Intensity) (e.g. “Morning”)</td>
</tr>
<tr>
<td></td>
<td>• (Location, Intensity) (e.g. “Living room”)</td>
</tr>
<tr>
<td></td>
<td>• (User, Intensity) (e.g. “Tom”)</td>
</tr>
<tr>
<td></td>
<td>• (Ambient state, Intensity) (e.g. odor, noise)</td>
</tr>
<tr>
<td></td>
<td>✓ Inside Features</td>
</tr>
<tr>
<td></td>
<td>• (Physical state, Intensity) (e.g. tiredness)</td>
</tr>
<tr>
<td></td>
<td>• (Psychological state, Intensity) (e.g. mood)</td>
</tr>
<tr>
<td></td>
<td>Problem Features</td>
</tr>
<tr>
<td></td>
<td>• (Motivational state, Intensity) (e.g. goal, initial state)</td>
</tr>
<tr>
<td></td>
<td>• (Outcome, Intensity) (e.g. success)</td>
</tr>
<tr>
<td>Sequential Features (Content)</td>
<td>Event 1 ➔ Event 2 ➔ ... ➔ Event n</td>
</tr>
<tr>
<td></td>
<td>Intensity ➔ Intensity ➔ ... ➔ Intensity</td>
</tr>
</tbody>
</table>
General Structure of Features

- Features
  - Repeatable
    - Sequential
      - Event
  - Unrepeatable
    - Parallel
      - Environment
      - Problem
        - Outside
        - Inside
Features in Measurement

• New Measurement
  – Activation factors
    • Frequency
    • Similarity (sequential vs. parallel)
    • Value of information (Importance)
    • Position (Primary effect & Recent effect)
    • Difficulty of understanding
      (ordering, concreteness, predictability, logical associations)
  – Forgetting factors (the same)
    • Times of retrieval
    • Time decay
Compute Intensity $I$

- **Context Similarity**
  - Given two context:
    $$cx_i = (f_{i_1}, f_{i_2}, \ldots, f_{i_n}), \quad cx_j = (f_{j_1}, f_{j_2}, \ldots, f_{j_n})$$
  - Compute Similarity
    $$r_{cx}(cx_i, cx_j) = \frac{2 \sum_{f \in cx_i \land cx_j} IA(f)}{\sum_{f \in cx_i} IA(f) + \sum_{f \in cx_j} IA(f)}$$

where $IA(\cdot)$ is the amount of information of a feature instance, which represents the importance of the instance in term of the feature. $cx_i \land cx_j$ represents the common feature instances of the two context. Those rarely happened features have more contribution to the similarity because they are more unique.
Compute Intensity I

• Content Similarity
  – Given two content:
    \[ cn_i^s = (ev_{i_1}^{(1)}, ev_{i_2}^{(2)}, \ldots, ev_{i_s}^{(s)}) \]
    \[ cn_j^t = (ev_{j_1}^{(1)}, ev_{j_2}^{(2)}, \ldots, ev_{j_t}^{(t)}) \]
  – Sequence Alignment[1]
    \[ \overline{cn_i^s} = (a_1, a_2, \ldots, a_l), a_i \in \{ev_{ik}^k\}_{k=1}^s \cup \{-\} \]
    \[ \overline{cn_j^t} = (b_1, b_2, \ldots, b_l), b_i \in \{ev_{jk}^k\}_{k=1}^t \cup \{-\} \]
  – Compute Similarity
    \[ r_{cn}(cn_i^s, cn_j^t) = \frac{2 \sum_{x \in \overline{cn_i^s} \cap \overline{cn_j^t}} IA(x)}{IA(cn_i^s) + IA(cn_j^t)} \]
    where \( \overline{cn_i^s} \cap \overline{cn_j^t} \) represents the set of common sequences

Compute Intensity $I$

- **General Amount of Information**

  \[
  Q(cx_i) = \sum_j p(cx_j) r_{cx}(cx_i, cx_j),
  \]
  \[
  Q(cn_i) = \sum_j p(cn_j) r_{cn}(cn_i, cn_j).
  \]

  \[
  \overline{IA}(cx_i) = -\log Q(cx_i),
  \]
  \[
  \overline{IA}(cn_i) = -\log Q(cn_i).
  \]

- **Memory Familiarity (MF)**

  \[
  F_M = \frac{1}{IA_M + 1} \in [0, 1],
  \]

- **Intensity**

  \[
  I^\text{once}_{cx} = F_{cx} \cdot V_{cx},
  \]
  \[
  I^\text{once}_{cn} = F_{cn} \cdot V_{cn}.
  \]
Compute Feature Intensity $I_i$

According to the Assumption 6 and 7, for each features, the once activation of the context or content distributes to its features based on their memory familiarity:

$$I_{once}^{f_i} = \frac{F_{f_i}}{\sum_{j=1}^{m} F_{f_j}} I_{once}^{c}$$, $i = 1, \ldots, m$,

$$I_{once}^{ev_i} = \frac{F_{ev_i|ev_{i-1}}}{\sum_{j=1}^{m} F_{ev_j|ev_{j-1}}} I_{once}^{c}$$, $i = 1, \ldots, n$. 

Update the total memory intensities of the features based on the principle that the more the total intensity near the upper bound 1, the less the total intensity increases:

$$I^t = I^{t-1} + I_{once}^t (1 - I^{t-1}).$$

Based on Ebbinghaus forgetting curve [11], the intensity of memory decay with time in an exponential way, and the speed of decay negatively relates with the initial memory intensity (see Figure 3).

$$I^t = I^{t-1} e^{-\frac{\delta \Delta t}{I^{t-1}}}.$$