Lecture 22

Two-layered Confabulation Architecture for Behavior Selection
Contents

1. Motivation of Research
2. Cogent Confabulation
3. Two-layered Confabulation Architecture
   1. Memory Module
   2. Behavior Module
      1. Will-based Confabulation
      2. Context-based Confabulation
      3. Arbiter
   3. Learning Module
4. Experimental Results
5. Conclusions
Motivation of Research

- Artificial creatures can be used as an intermediate interface for interactions between humans and service robots.

- Artificial creatures need to imitate real creatures for more natural interaction with human beings
  - By selecting various proper behaviors referring to its internal state and the context in the environment

- Confabulation process for their behavior selection is based on the Mechanism of Thought.
Motivation of Research

Typical control architecture

- A behavior is selected in behavior module considering the perception and motivation.
Motivation of Research

- Priority-based behavior selector
  - Too deterministic to realize a deliberative behavior

```
Context-based Behavior

Selected Behavior

Internal State-based Behavior
```

Confabulations, where millions of items of relevant knowledge are applied in parallel in the human brain, are typically employed in thinking.

Knowledge links are formed between meaningfully co-occurring symbols, essentially as postulated by Hebb.
Probabilistic reasoning
  - Assumed fact symbols: \(a, b, c, d\)
    - Representing contexts and internal states
  - Conclusion symbol: \(e\)
    - Behavior

Confidence to conclusion in Bayesian reasoning
  - Posterior probability: \(p(e \mid abcd)\)
  - Difficult to build the probability database

Confidence to conclusion in cogent confabulation
  - Cogency: \(p(abcd \mid e)\)
  - Easier to build the probability database
Cogency:

\[ p(abcd \mid e) \]

\[
p(abcd \mid e)^4 = \left[ p(abcde) / p(ae) \right] \cdot \left[ p(abcde) / p(be) \right] \cdot \left[ p(abcde) / p(ce) \right] \cdot \left[ p(abcde) / p(de) \right] \cdot \left[ p(a \mid e) \cdot p(b \mid e) \cdot p(c \mid e) \cdot p(d \mid e) \right]
\]

\[
p(abcd \mid e)^4 \approx k \cdot \left[ p(a \mid e) \cdot p(b \mid e) \cdot p(c \mid e) \cdot p(d \mid e) \right]
\]

Having the maximum cogency is approximately equivalent to maximizing the following probability:

\[ p(a \mid e) \cdot p(b \mid e) \cdot p(c \mid e) \cdot p(d \mid e) \]
Two-layered Confabulation

- Perception Module
  - Context Module
  - Internal State Module
  - Learning Module
  - Memory Module

- External Environment
  - Behavior Module
    - Context-based Confabulation
    - Will-based Confabulation
  - Arbiter
  - Actuator Module
Two-layered Confabulation

- Conditional probabilities between behaviors and each of the internal states and contexts are stored in memory.
Two-layered Confabulation
Two-layered Confabulation

Behavior Module

Context-based Confabulation

Will-based Confabulation

Recommended Behavior 1
Recommended Behavior 2
Recommended Behavior 3
...

Arbiter
Two-layered Confabulation

Behavior Module

Context-based Confabulation

Recommended Behavior 1

Recommended Behavior 2

Recommended Behavior 3

Will-based Confabulation

Context-based Cogency

Context-based Cogency

Context-based Cogency

... Recommend Behavior

... Recommend Behavior

... Recommend Behavior

Will-based Cogency

Will-based Cogency

Will-based Cogency

... Recommend Behavior

... Recommend Behavior

... Recommend Behavior

Arbiter

Finally Selected Behavior
Will-based Confabulation

- Cogency values between behaviors and current wills are calculated as follows:

\[
E_{\text{will}}(b_i) = p(w_1 | b_i) \cdot p(w_2 | b_i) \cdot \cdots \cdot p(w_m | b_i)
\]

- Some behaviors with high cogency values are recommended to the next context-based confabulation sub-module.
Context-based Confabulation

\[
E_{\text{context}}(b_r^1) = p(c_1 \mid b_r^1) \cdot p(c_2 \mid b_r^1) \cdots p(c_n \mid b_r^1)
\]
\[
E_{\text{context}}(b_r^2) = p(c_1 \mid b_r^2) \cdot p(c_2 \mid b_r^2) \cdots p(c_n \mid b_r^2)
\]
\[\vdots\]
\[
E_{\text{context}}(b_r^k) = p(c_1 \mid b_r^k) \cdot p(c_2 \mid b_r^k) \cdots p(c_n \mid b_r^k)
\]

\(E_{\text{context}}\): cogency value (excitation) of the behavior to the context
\(c_i\): the \(i\)-th context
\(n\): the number of considered contexts
\(b_r^j\): the \(j\)-th recommended behavior from the will-based confabulation
\(k\): the number of recommended behaviors

\(p(c_i \mid b_r^j)\): conditional probability between the \(i\)-th context and the \(j\)-th recommended behavior
Arbiter

To decide the final behavior among $k$ candidates based on the results of will-based confabulation $E_{\text{will}}(b_j)$ and context-based confabulation $E_{\text{context}}(b^r_j)$.

Behavior is determined by the max-product operation:

$$E_{\text{arbiter}}(b^s) = \max_{j} \left[ E_{\text{will}}(b^r_j) \cdot E_{\text{context}}(b^r_j) \right], \quad j = 1, \ldots, k$$

where $b^s$ is the finally selected behavior.
## Memory Module

<table>
<thead>
<tr>
<th></th>
<th>FOOD1</th>
<th>FOOD2</th>
<th>BED1</th>
<th>TOILET1</th>
<th>TOY1</th>
<th>TOY2</th>
<th>TOY3</th>
<th>HUMAN1</th>
<th>COMRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_STOP</td>
<td>0.052867</td>
<td>0.066835</td>
<td>0.030952</td>
<td>0.063683</td>
<td>0.041612</td>
<td>0.031723</td>
<td>0.063462</td>
<td>0.077256</td>
<td>0.010584</td>
</tr>
<tr>
<td>B_SIT</td>
<td>0.054406</td>
<td>0.021737</td>
<td>0.079182</td>
<td>0.031868</td>
<td>0.030904</td>
<td>0.016136</td>
<td>0.047091</td>
<td>0.056464</td>
<td>0.033554</td>
</tr>
<tr>
<td>B_CROUCH</td>
<td>0.017003</td>
<td>0.068999</td>
<td>0.069999</td>
<td>0.053179</td>
<td>0.087168</td>
<td>0.040631</td>
<td>0.019635</td>
<td>0.029379</td>
<td>0.021205</td>
</tr>
<tr>
<td>B_SHAKE_H</td>
<td>0.012034</td>
<td>0.035317</td>
<td>0.064845</td>
<td>0.069946</td>
<td>0.054826</td>
<td>0.059851</td>
<td>0.010939</td>
<td>0.057509</td>
<td>0.049257</td>
</tr>
<tr>
<td>B_LIFT_ARM</td>
<td>0.07283</td>
<td>0.030277</td>
<td>0.008674</td>
<td>0.050849</td>
<td>0.029531</td>
<td>0.073517</td>
<td>0.039336</td>
<td>0.49722</td>
<td>0.065783</td>
</tr>
<tr>
<td>B_WHINE</td>
<td>0.061178</td>
<td>0.078347</td>
<td>0.05132</td>
<td>0.003339</td>
<td>0.027057</td>
<td>0.073524</td>
<td>0.06398</td>
<td>0.004303</td>
<td>0.010069</td>
</tr>
<tr>
<td>B_GO_BACK</td>
<td>0.027052</td>
<td>0.004182</td>
<td>0.010016</td>
<td>0.06143</td>
<td>0.030592</td>
<td>0.000536</td>
<td>0.035793</td>
<td>0.07166</td>
<td>0.005316</td>
</tr>
<tr>
<td>B_STEP_BK</td>
<td>0.078937</td>
<td>0.085721</td>
<td>0.086771</td>
<td>0.087075</td>
<td>0.080818</td>
<td>0.077174</td>
<td>0.085034</td>
<td>0.078362</td>
<td>0.084948</td>
</tr>
<tr>
<td>B_STEP_BT</td>
<td>0.074334</td>
<td>0.075793</td>
<td>0.08713</td>
<td>0.087421</td>
<td>0.080424</td>
<td>0.082993</td>
<td>0.061543</td>
<td>0.076015</td>
<td>0.07888</td>
</tr>
<tr>
<td>B_WANDER</td>
<td>0.022151</td>
<td>0.042755</td>
<td>0.011876</td>
<td>0.05866</td>
<td>0.049437</td>
<td>0.04645</td>
<td>0.08424</td>
<td>0.000616</td>
<td>0.022224</td>
</tr>
<tr>
<td>B_WANDER_F</td>
<td>0.015641</td>
<td>0.029322</td>
<td>0.04352</td>
<td>0.007527</td>
<td>0.056146</td>
<td>0.033274</td>
<td>0.026918</td>
<td>0.01072</td>
<td>0.00000</td>
</tr>
<tr>
<td>B_MOVE_C</td>
<td>0.056464</td>
<td>0.02714</td>
<td>0.005312</td>
<td>0.009975</td>
<td>0.052794</td>
<td>0.037931</td>
<td>0.054944</td>
<td>0.052168</td>
<td>0.00781</td>
</tr>
<tr>
<td>B_MOVE_C</td>
<td>0.040588</td>
<td>0.02347</td>
<td>0.025675</td>
<td>0.043569</td>
<td>0.019517</td>
<td>0.057332</td>
<td>0.028304</td>
<td>0.025319</td>
<td>0.025349</td>
</tr>
<tr>
<td>B_MOVE_S</td>
<td>0.031449</td>
<td>0.050965</td>
<td>0.030715</td>
<td>0.006897</td>
<td>0.031452</td>
<td>0.033686</td>
<td>0.012337</td>
<td>0.022519</td>
<td>0.053828</td>
</tr>
<tr>
<td>B_FOLLOW</td>
<td>0.008005</td>
<td>0.040605</td>
<td>0.013496</td>
<td>0.01263</td>
<td>0.016395</td>
<td>0.006428</td>
<td>0.014587</td>
<td>0.191563</td>
<td>0.190647</td>
</tr>
<tr>
<td>B_FOLLOW_F</td>
<td>0.015135</td>
<td>0.002878</td>
<td>0.007609</td>
<td>0.010402</td>
<td>0.015089</td>
<td>0.014214</td>
<td>0.016939</td>
<td>0.163516</td>
<td>0.174813</td>
</tr>
<tr>
<td>B_APPROACH</td>
<td>0.021015</td>
<td>0.001918</td>
<td>0.036216</td>
<td>0.042422</td>
<td>0.031572</td>
<td>0.007508</td>
<td>0.021871</td>
<td>0.751858</td>
<td>0.004476</td>
</tr>
<tr>
<td>B_MOVE_A</td>
<td>0.007681</td>
<td>0.01258</td>
<td>0.008477</td>
<td>0.001713</td>
<td>0.015815</td>
<td>0.005216</td>
<td>0.015323</td>
<td>0.185777</td>
<td>0.184116</td>
</tr>
<tr>
<td>B_MOVE_A</td>
<td>0.000896</td>
<td>0.015614</td>
<td>0.015561</td>
<td>0.003852</td>
<td>0.014852</td>
<td>0.007293</td>
<td>0.014877</td>
<td>0.193394</td>
<td>0.179043</td>
</tr>
<tr>
<td>B_SNIFE_H</td>
<td>0.012071</td>
<td>0.017292</td>
<td>0.005638</td>
<td>0.000177</td>
<td>0.002631</td>
<td>0.000981</td>
<td>0.000378</td>
<td>0.190822</td>
<td>0.194179</td>
</tr>
<tr>
<td>B_LOOK_A</td>
<td>0.070174</td>
<td>0.072039</td>
<td>0.033595</td>
<td>0.025405</td>
<td>0.045654</td>
<td>0.069356</td>
<td>0.033961</td>
<td>0.049734</td>
<td>0.000436</td>
</tr>
<tr>
<td>B_LOOK_A</td>
<td>0.049158</td>
<td>0.050885</td>
<td>0.048245</td>
<td>0.054141</td>
<td>0.096236</td>
<td>0.100986</td>
<td>0.100541</td>
<td>0.098318</td>
<td>0.098225</td>
</tr>
<tr>
<td>B_EAT_CHE</td>
<td>0.402917</td>
<td>0.427327</td>
<td>0.018855</td>
<td>0.010522</td>
<td>0.036426</td>
<td>0.010731</td>
<td>0.011324</td>
<td>0.265207</td>
<td>0.001379</td>
</tr>
<tr>
<td>B_EAT_SLC</td>
<td>0.419448</td>
<td>0.419582</td>
<td>0.01139</td>
<td>0.004073</td>
<td>0.005075</td>
<td>0.000981</td>
<td>0.009821</td>
<td>0.153668</td>
<td>0.037569</td>
</tr>
</tbody>
</table>
Learning Module

- All the memory contents are provided by an expert as an initialization process on plausible behaviors of the artificial creature.

- As it should continuously adapt to the varying environment or to the user’s preference, the memory contents should be updated in the learning module.

- Reinforcement learning as a training process is provided like a real pet training.

- The learning module updates the memory contents according to the user-given reward or penalty signal.
Learning Module

- The user grants either reward or penalty by patting or hitting the artificial creature to teach a desired behavior at a given situation.

\[
p_{\text{temp}}(c_i | b_j) = \begin{cases} 
    p_t(c_i | b_j) + \delta & \text{ (if input is a reward) } \\
    p_t(c_i | b_j) - \delta & \text{ (if input is a penalty) } \\
    p_t(c_i | b_j) & \text{ (otherwise) }
\end{cases}
\]

- As the sum of all probabilities must be equal to one, the following normalization process is applied:

\[
p_{t+1}(c_i | b_j) = \frac{p_{\text{temp}}(c_i | b_j)}{n \sum_{i=1}^{n} p_{\text{temp}}(c_i | b_j)} 
\]

- \( b_j \): selected behavior
- \( p_t \): behavior probability at time \( t \)
- \( \delta \): learning amount
- \( c_i \): context
- \( n \): the number of contexts
- \( c_i \): the \( i \)-th context
Experimental Results
### Experimental Results

<table>
<thead>
<tr>
<th>Internal State</th>
<th>Assumed fact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation</td>
<td>Play</td>
</tr>
<tr>
<td></td>
<td>Comradeship</td>
</tr>
<tr>
<td></td>
<td>Shelter-Seeking</td>
</tr>
<tr>
<td></td>
<td>Pain</td>
</tr>
<tr>
<td>Homeostasis</td>
<td>Energy</td>
</tr>
<tr>
<td></td>
<td>Fatigue</td>
</tr>
<tr>
<td></td>
<td>Excretion</td>
</tr>
<tr>
<td>Emotion</td>
<td>Happy</td>
</tr>
<tr>
<td></td>
<td>Sad</td>
</tr>
<tr>
<td></td>
<td>Anger</td>
</tr>
<tr>
<td></td>
<td>Fear</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context</th>
<th>Assumed fact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Morning</td>
</tr>
<tr>
<td></td>
<td>Afternoon</td>
</tr>
<tr>
<td></td>
<td>Evening</td>
</tr>
<tr>
<td></td>
<td>Night</td>
</tr>
<tr>
<td>Place</td>
<td>Bed Room</td>
</tr>
<tr>
<td></td>
<td>Toilet</td>
</tr>
<tr>
<td>Object</td>
<td>Food</td>
</tr>
<tr>
<td></td>
<td>Comrade</td>
</tr>
<tr>
<td></td>
<td>Toy</td>
</tr>
</tbody>
</table>
Since the user forbade eating behavior around the toilet, the probability of Rity’s eating there became lower.
The behavior, which is not related to the context, would be selected if Rity has a strong will.
Conclusion

- Two-layered confabulation architecture for behavior selection referring to the contexts and internal states.
- In the will-based confabulation, behaviors are selected considering the internal states such as motivation, homeostasis and emotions.
- The selected behaviors are forwarded to the context-based confabulation to consider the contexts on “when,” “where” and “what.”
- The arbiter finally decides the most proper behavior among the suggested behaviors.
- A learning mechanism mimicking the real pet training was employed.