Attention-based Dialog State Tracking for Conversational Interview Coaching

Ming-Hsiang Su, Chung-Hsien Wu, Kun-Yi Huang and Chu-Kwang Chen

Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, TAIWAN

Introduction

- **Goal:** Propose an attention weighting mechanism for dialog state tracking in a conversational interview coaching system.
- **Problem:**
  1. The **semantic slots** for dialog state representation are difficult to define manually.
  2. The interviewee’s response contains **irrelevant sentences**.
- **Approach:**
  1. **Topic profile similarity** between question and a response sentence is estimated based on LDA-based topic model.
  2. **Relevance score** between question and a response sentence is obtained based on convolutional neural tensor network (CNTN).
  3. **Attention mechanism** for sentence weighting.
  4. **Two-layer LSTM-based autoencoder** for dialog state tracking.

Experimental Results

- **NCKU Interview Database** collected 260 dialogs with 1754 ordinary questions and 1262 follow-up questions.

<table>
<thead>
<tr>
<th>Total</th>
<th>Number of turns</th>
<th>Average number of turns</th>
<th>Average number of sentences in each answer</th>
<th>Interview time (minutes) per interview</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3016</td>
<td>10.7</td>
<td>3.84</td>
<td>20</td>
</tr>
</tbody>
</table>

- The accuracy of CNTN-based attention model achieved **89.87%**, when the tensor dimensionality was 1 and the number of CNN filters was 128.

- **LSTM-based autoencoder** was evaluated by visualizing the vector representation of input and output sequence, when the number of topics was 30.

Performance Comparison between the traditional and the proposed methods using the same Double Q-learning agent.

Conclusions

- A five-fold cross validation scheme was employed and the results show that the proposed method outperformed the semantic slot-based method.