A Network-based End-to-End Trainable Task-oriented Dialogue System

Authors: Tsung-Hsien Wen, David Vandyke, Nikola Mrkšić, Milica Gašić, Lina M. Rojas-Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young

Presented by: Qihao Shao
Overview

• Introduction
• Model
• Wizard-of-Oz Data Collection
• Empirical Experiments
• Conclusions
Overview

• Introduction
• Model
• Wizard-of-Oz Data Collection
• Empirical Experiments
• Conclusions
Introduction

• Treat as a POMDP and use RL to train dialogue policies
• Build end-to-end trainable, non-task-oriented conversational systems using seq2seq model
• The authors propose a model by balancing the strengths and the weaknesses of these two
Overview

• Introduction
• Model
• Wizard-of-Oz Data Collection
• Empirical Experiments
• Conclusions
Model

- Intent Network
- Belief Trackers
- Database Operator
- Policy Network
- Generation Network
Model
Intent Network

• Encoder in the sequence-to-sequence framework
• Typically, an LSTM network is used

\[ z_t = z_t^N = \text{LSTM}(w_0^t, w_1^t, \ldots w_N^t) \]

• Alternatively, a CNN can be used

\[ z_t = \text{CNN}(w_0^t, w_1^t, \ldots w_N^t) \]
Intent Network
Belief Trackers

• Core component of the model
• Every slot has its belief tracker
• Each tracker is a Jordan type RNN with a CNN feature extractor
Belief Trackers

\[ f_{v,cnn}^t = \text{CNN}_{s,v}^{(u)}(u_t) \oplus \text{CNN}_{s,v}^{(m)}(m_{t-1}) \]

\[ f_v^t = f_{v,cnn}^t \oplus p^{t-1}_v \oplus p^t_{\emptyset} \]

\[ g_v^t = w_s \cdot \text{sigmoid}(W_s f_v^t + b_s) + b'_s \]

\[ p_v^t = \frac{\exp(g_v^t)}{\exp(g_{\emptyset,s}) + \sum_{v' \in V_s} \exp(g_{v'}^t)} \]
Belief Trackers
Belief Trackers
Database Operator

• The DB query $q_t$ is formed by

$$q_t = \bigcup_{s' \in S_I} \{ \arg\max_{v} p_{s'}^t \}$$

• Then query is applied to the DB to create a binary truth value vector $x_t$ over DB entities

• The entity referenced by the entity pointer is used to form the final system response
Database Operator

Intent Network

Can I have <v.food>

Generation Network

<v.name> serves great <v.food>

Policy Network

Korean 0.7
British 0.2
French 0.1

MySQL query: "Select * where food=Korean"

Belief Tracker

Can I have <v.food>

Database Operator

Copy field

DB pointer

0 0 0 ... 0 1
Policy Network

• Can be viewed as the glue binding other modules together

\[ o_t = \tanh(W_{zo}z_t + W_{po}\hat{p}_t + W_{xo}\hat{x}_t) \]
Policy Network
Generation Network

\[ P(w_{j+1}^t | w_j^t, h_{j-1}^t, o_t) = \text{LSTM}_j(w_j^t, h_{j-1}^t, o_t) \]

• Once the output token sequence has been generated, the generic tokens are replaced by their actual values.
Generation Network

• Attentive Generation Network

\[
o_{t}^{(j)} = \tanh(W_{zo}z_{t} + \hat{p}_{t}^{(j)} + W_{xo}\hat{x}_{t})
\]

\[
\hat{p}_{t}^{(j)} = \sum \alpha_{s}^{(j)} \tanh(W_{po}^{s} \cdot \hat{p}_{s}^{t})
\]

\[
\alpha_{s}^{(j)} = \text{softmax} (r^{T} \tanh(W_{r} \cdot u_{t}))
\]

\[
u_{t} = z_{t} \oplus \hat{x}_{t} \oplus \hat{p}_{s}^{t} \oplus w_{j}^{t} \oplus h_{j-1}^{t}
\]
Overview

• Introduction
• Model
• **Wizard-of-Oz Data Collection**
• Empirical Experiments
• Conclusions
Wizard-of-Oz Data Collection

• This paper proposed a novel crowdsourcing version of the Wizard-of-Oz paradigm

• Designed two webpages on Amazon Mechanical Turk, one for wizards and the other for users
Wizard-of-Oz Data Collection

Task 02004: You are looking for and it should serve gastropub food. You don't care about the price range. You want to know the address.

Info Desk: Hello, welcome to the Cambridge restaurant system. You can ask for restaurants by area, price range or food type. How may I help you?
Customer: I want a gastropub food
Info Desk: There are 4 restaurants serving gastropub food, what price range do you want?

[Dialogue continues with options for price range]

Customer: I don't care about the price range, just give me the address please.

[Submit the HIT button]
Wizard-of-Oz Data Collection
Wizard-of-Oz Data Collection

• 99 restaurants in the DB
• 3000 HITs (Human Intelligence Tasks) in total
• 680 dialogues after data cleaning
• Cost ~ 400 USD
Overview

• Introduction
• Model
• Wizard-of-Oz Data Collection
• Empirical Experiments
• Conclusions
## Empirical Experiments

<table>
<thead>
<tr>
<th>Tracker type</th>
<th>Informable</th>
<th></th>
<th>Requestable</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prec.</td>
<td>Recall</td>
<td>F-1</td>
<td>Prec.</td>
</tr>
<tr>
<td>cnn</td>
<td>99.77%</td>
<td>96.09%</td>
<td>97.89%</td>
<td>98.66%</td>
</tr>
<tr>
<td>ngram</td>
<td>99.34%</td>
<td>94.42%</td>
<td>96.82%</td>
<td>98.56%</td>
</tr>
</tbody>
</table>
## Empirical Experiments

<table>
<thead>
<tr>
<th>Encoder</th>
<th>Tracker</th>
<th>Decoder</th>
<th>Match(%)</th>
<th>Success(%)</th>
<th>T5-BLEU</th>
<th>T1-BLEU</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lstm</td>
<td>-</td>
<td>lstm</td>
<td>-</td>
<td>-</td>
<td>0.1650</td>
<td>0.1718</td>
</tr>
<tr>
<td>lstm</td>
<td>turn recurrence</td>
<td>lstm</td>
<td>-</td>
<td>-</td>
<td>0.1813</td>
<td>0.1861</td>
</tr>
<tr>
<td><strong>Variant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lstm</td>
<td>rnn-cnn, w/o req.</td>
<td>lstm</td>
<td>89.70</td>
<td>30.60</td>
<td>0.1769</td>
<td>0.1799</td>
</tr>
<tr>
<td>cnn</td>
<td>rnn-cnn</td>
<td>lstm</td>
<td>88.82</td>
<td>58.52</td>
<td>0.2354</td>
<td>0.2429</td>
</tr>
<tr>
<td><strong>Full model w/ different decoding strategy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lstm</td>
<td>rnn-cnn</td>
<td>lstm</td>
<td>86.34</td>
<td>75.16</td>
<td>0.2184</td>
<td>0.2313</td>
</tr>
<tr>
<td>lstm</td>
<td>rnn-cnn + weighted</td>
<td></td>
<td>86.04</td>
<td>78.40</td>
<td>0.2222</td>
<td>0.2280</td>
</tr>
<tr>
<td>lstm</td>
<td>rnn-cnn + att.</td>
<td></td>
<td></td>
<td></td>
<td>0.2286</td>
<td>0.2388</td>
</tr>
<tr>
<td>lstm</td>
<td>rnn-cnn + att. + weighted</td>
<td></td>
<td>90.88</td>
<td>83.82</td>
<td>0.2304</td>
<td>0.2369</td>
</tr>
</tbody>
</table>
## Empirical Experiments

<table>
<thead>
<tr>
<th>Metric</th>
<th>NN</th>
<th>HDC</th>
<th>Tie</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Success</strong></td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comprehension</strong></td>
<td>4.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Naturalness</strong></td>
<td>4.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subj. Success</strong></td>
<td>96.95%</td>
<td>95.12%</td>
<td>-</td>
</tr>
<tr>
<td><strong>Avg. # of Turn</strong></td>
<td>3.95</td>
<td>4.54</td>
<td>-</td>
</tr>
</tbody>
</table>

### Comparisons (%)

- Naturalness: 46.95% (p < 0.005)
- Comprehension: 45.12% (p < 0.005)
- Preference: 50.00%
- Performance: 43.90%

* p < 0.005, # of comparisons: 164
Overview

• Introduction
• Model
• Wizard-of-Oz Data Collection
• Empirical Experiments
• Conclusions
Conclusions

• Combines modularly connected model and end-to-end trainable model
• First end-to-end NN-based model that can conduct meaningful dialogues in a task-oriented application
Thank you