N-Shot Learning for Augmenting Task-Oriented Dialogue State Tracking

Taha Aksu¹,², Zhengyuan Liu²,³, Min-Yen Kan¹, and Nancy F.Chen²

¹ Web IR / NLP Group (WING), National University of Singapore
² Institute for Infocomm Research, A*STAR
³ CNRS@CREATE

ACL, 22-27 May 2022
Introduction

Dialogue State Tracking (DST)

- Extract key attributes of users preferences.

Low Resource Domain Adaptation in DST

- Crucial because: New business models $\rightarrow$ new DST agents in new domains.
- Challenging because: New domains $\rightarrow$ new slot labels and values with less data.

Data Augmentation for TODs

- Augmentation is good to address lack of data.
- TODs harbor even more potential for augmentation.

---

U0: I would like a British food restaurant in the centre.
BS: [restaurant-food: British]

S1: Sure, there are 7 restaurants that meet your needs. 4 are moderate and 3 are expensive. Do you have a preference?

U1: Only the best for my family .. we'll take the expensive one. Book us a table for 5 at 14:00 on Thursday.
BS: [restaurant-food: British, restaurant-pricerange: expensive, restaurant-bookpeople: 5, restaurant-booktime: 14:00, restaurant-bookday: Thursday]

S2: I'm sorry I am having difficulty making a reservation for you. Shall we try another time or restaurant type?

U2: Let's try Italian instead.
BS: [restaurant-food: Italian]
Dialogue A

U → Hi, I am looking for a train that is going to Cambridge and arriving there by 20:45, is there anything like that?
A → There are many trains like that. Where will you be departing from?
U → I am departing from Birmingham New Street.
A → Can you confirm your desired travel day?
U → I would like to leave on Wednesday.

Dialogue B

U → I need a trip going to Cambridge and arrive by 11:00.
A → Where are you leaving from?
U → I am leaving from Norwich.
A → We have many trains that meet that criteria. Can you tell me what day you are looking to travel on?
U → Of course, I will be traveling on Tuesday.

New Dialogue

U → I need a trip going to Cambridge and arrive by 11:00.
A → There are many trains like that. Where will you be departing from?
U → I am departing from Norwich.
A → We have many trains that meet that criteria. Can you tell me what day you are looking to travel on?
U → Of course, I will be traveling on Tuesday.
U → Hi, I am looking for a train that is going to Cambridge and arriving there by 20:45, is there anything like that?
A → There are many trains like that. Where will you be departing from?
U → I am departing from Birmingham New Street.
A → Can you confirm your desired travel day?
U → I would like to leave on Wednesday.
A → Okay, we have a ticket that is fit, should I book it?
U → Yes, please.
### Experiments and Results

MultiWOZ dataset – with 5 domains over 10k dialogues.

- Leave one domain out → **train** on rest then **fine tune** and **test** on the left out domain.
- During **fine tuning** use 5/10 shots or augmented dialogues from the same shots.
- We report results with both **TRADE** and **TOD-BERT**.

<table>
<thead>
<tr>
<th></th>
<th>Hotel Joint</th>
<th>Hotel Slot</th>
<th>Taxi Joint</th>
<th>Taxi Slot</th>
<th>Restaurant Joint</th>
<th>Restaurant Slot</th>
<th>Attraction Joint</th>
<th>Attraction Slot</th>
<th>Train Joint</th>
<th>Train Slot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Model (BM) trained on other 4 domains</td>
<td>0.12</td>
<td>0.64</td>
<td>0.60</td>
<td>0.73</td>
<td>0.12</td>
<td>0.54</td>
<td>0.18</td>
<td>0.54</td>
<td>0.22</td>
<td>0.49</td>
</tr>
<tr>
<td>BM fine tuned with 1% data (84 samples)</td>
<td>0.21</td>
<td>0.76</td>
<td>0.61</td>
<td>0.75</td>
<td>0.21</td>
<td>0.77</td>
<td>0.43</td>
<td>0.74</td>
<td>0.61</td>
<td>0.91</td>
</tr>
<tr>
<td>5-Shot Augmentation on Target Domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM fine-tuned with 5 samples</td>
<td>0.12</td>
<td>0.65</td>
<td>0.59</td>
<td>0.75</td>
<td>0.12</td>
<td>0.58</td>
<td>0.25</td>
<td>0.59</td>
<td>0.25</td>
<td>0.66</td>
</tr>
<tr>
<td>BM fine-tuned with augmented samples</td>
<td>0.12</td>
<td><strong>0.67</strong>*</td>
<td>0.58</td>
<td>0.75</td>
<td><strong>0.13</strong>*</td>
<td><strong>0.62</strong>*</td>
<td><strong>0.26</strong>*</td>
<td><strong>0.61</strong></td>
<td><strong>0.31</strong>*</td>
<td><strong>0.77</strong>*</td>
</tr>
<tr>
<td>10-Shot Augmentation on Target Domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM fine-tuned with 10 samples</td>
<td>0.14</td>
<td>0.68</td>
<td>0.60</td>
<td><strong>0.76</strong></td>
<td>0.13</td>
<td>0.63</td>
<td>0.30</td>
<td>0.63</td>
<td>0.37</td>
<td>0.81</td>
</tr>
<tr>
<td>BM fine-tuned with augmented samples</td>
<td><strong>0.15</strong></td>
<td><strong>0.69</strong></td>
<td>0.60</td>
<td><strong>0.76</strong></td>
<td><strong>0.16</strong>*</td>
<td><strong>0.70</strong>*</td>
<td><strong>0.32</strong>*</td>
<td><strong>0.66</strong>*</td>
<td><strong>0.39</strong></td>
<td><strong>0.83</strong></td>
</tr>
</tbody>
</table>
Experiments and Results

MultiWOZ dataset – with 5 domains over 10k dialogues.
- Leave one domain out → **train** on rest then **fine tune** and **test** on the left out domain.
- During **fine tuning** use 5/10 shots or augmented dialogues from the same shots.
- We report results with both **TRADE** and TOD-BERT.

CoCo → learning based augmentation model.
- CoCo: Trained on the whole training set of MultiWOZ.
- Our model utilizes only the shots provided.

<table>
<thead>
<tr>
<th>5 Shot Augmentation on Target Domain</th>
<th>Hotel Joint</th>
<th>Hotel Slot</th>
<th>Taxi Joint</th>
<th>Taxi Slot</th>
<th>Restaurant Joint</th>
<th>Restaurant Slot</th>
<th>Attraction Joint</th>
<th>Attraction Slot</th>
<th>Train Joint</th>
<th>Train Slot</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM fine-tuned with CoCo</td>
<td>0.12</td>
<td>0.66</td>
<td><strong>0.60</strong></td>
<td>0.75</td>
<td>0.13</td>
<td>0.62</td>
<td>0.24</td>
<td>0.58</td>
<td>0.27</td>
<td>0.69</td>
</tr>
<tr>
<td>BM fine-tuned with our framework</td>
<td>0.12</td>
<td><strong>0.67</strong></td>
<td>0.58</td>
<td>0.75</td>
<td>0.13</td>
<td>0.62</td>
<td>0.26</td>
<td>0.61</td>
<td><strong>0.31</strong></td>
<td><strong>0.77</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10 Shot Augmentation on Target Domain</th>
<th>Hotel Joint</th>
<th>Hotel Slot</th>
<th>Taxi Joint</th>
<th>Taxi Slot</th>
<th>Restaurant Joint</th>
<th>Restaurant Slot</th>
<th>Attraction Joint</th>
<th>Attraction Slot</th>
<th>Train Joint</th>
<th>Train Slot</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM fine-tuned with CoCo</td>
<td><strong>0.15</strong></td>
<td>0.68</td>
<td><strong>0.61</strong></td>
<td>0.75</td>
<td><strong>0.16</strong></td>
<td>0.67</td>
<td>0.31</td>
<td>0.64</td>
<td><strong>0.39</strong></td>
<td><strong>0.82</strong></td>
</tr>
<tr>
<td>BM fine-tuned with our framework</td>
<td><strong>0.15</strong></td>
<td><strong>0.69</strong></td>
<td><strong>0.60</strong></td>
<td><strong>0.76</strong></td>
<td><strong>0.16</strong></td>
<td><strong>0.70</strong></td>
<td><strong>0.32</strong></td>
<td><strong>0.66</strong></td>
<td><strong>0.39</strong></td>
<td><strong>0.83</strong></td>
</tr>
</tbody>
</table>
Conclusion

- TOD Augmentation framework
  - dataset level modifications.
  - rather than on datum/sample level.
- Organized structure in a TODs belief state is an effective way to
  - assign functions to turn pairs.
  - break down dialogues to pieces.
- Augmentation ✅
  - Intent recognition ❓
  - Response generation ❓

Thanks for listening!
Please reach out to me for further discussion and questions:
- Website: https://cuthalionn.github.io/
- Email: taksu@u.nus.edu