Institute for
Infocomm Research

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

Taha Aksu[1,2], Zhengyuan Liu[2,3], Min-Yen Kan[1], and Nancy F.Chen[2]

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 → new DST agents in new domains.
- **Challenging** because: New domains → new slot labels and values with less data.

Data Augmentation for TODs

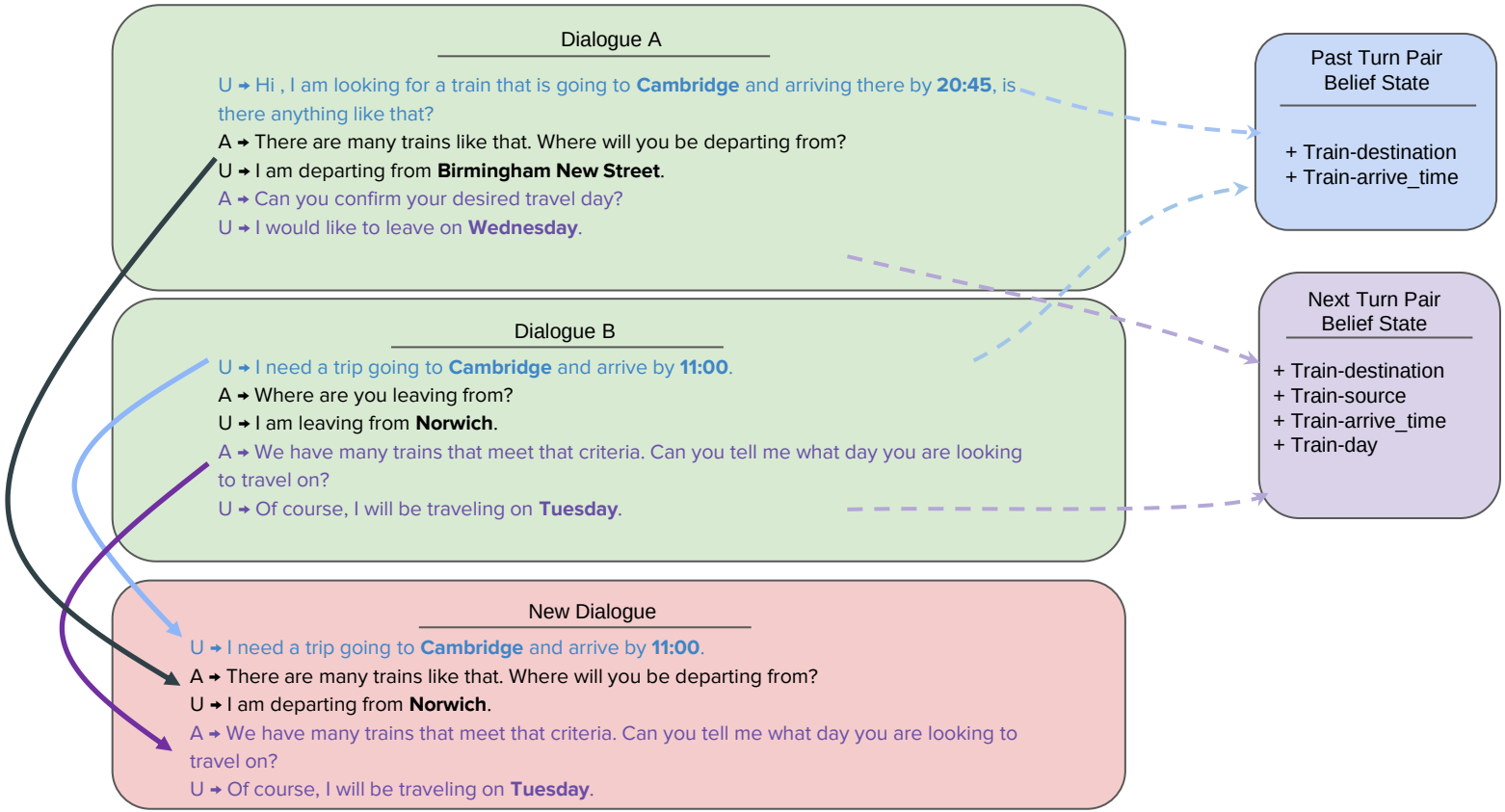
- Augmentation is good to address lack of data.
- TODs harbor even more potential for augmentation.

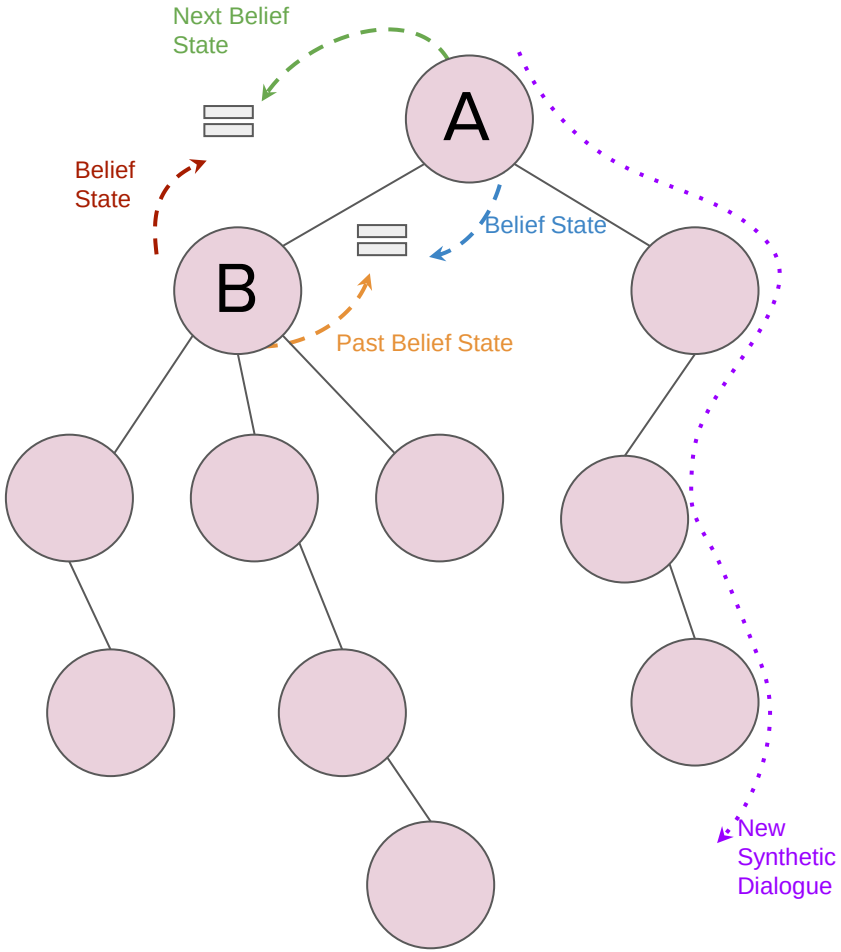
U0: I would like a British food restaurant in the centre.
BS: {restaurant-food: **British**}

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**}

Intuition


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.
Past Turn Pair Belief State
+ Train-destination
+ Train-arrive_time
Next Turn Pair Belief State
+ Train-destination
+ Train-source
+ Train-arrive_time
+ Train-day



Original Dialogue

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.

Turn-pair template

De-lexicalized Turns:
A → There are many trains like that . where will you be departing from?
U → I am departing from **[train-departure]**.

Turn-pair Function

BS: {train-destination, train-arrive_time, train-departure}
Past BS: {train-destination, train-arrive_time}
Next BS:{train-destination, train-arrive_time, train-departure, train-day}

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.

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

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.

	Hotel		Taxi		Restaurant		Attraction		Train	
	Joint	Slot	Joint	Slot	Joint	Slot	Joint	Slot	Joint	Slot
5 Shot Augmentation on Target Domain										
BM fine-tuned with CoCo	0.12	0.66	**0.60**	0.75	**0.13**	**0.62**	0.24	0.58	0.27	0.69
BM fine-tuned with our framework	0.12	**0.67**	0.58	0.75	**0.13**	**0.62**	**0.26**	**0.61**	**0.31**	**0.77**
10 Shot Augmentation on Target Domain										
BM fine-tuned with CoCo	**0.15**	0.68	**0.61**	0.75	**0.16**	0.67	0.31	0.64	**0.39**	0.82
BM fine-tuned with our framework	**0.15**	**0.69**	0.60	**0.76**	**0.16**	**0.70**	**0.32**	**0.66**	**0.39**	**0.83**

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

Scan to read the paper!