Memory Attention Neural Network for Multi-Domain Dialogue State Tracking

Zihan Xu

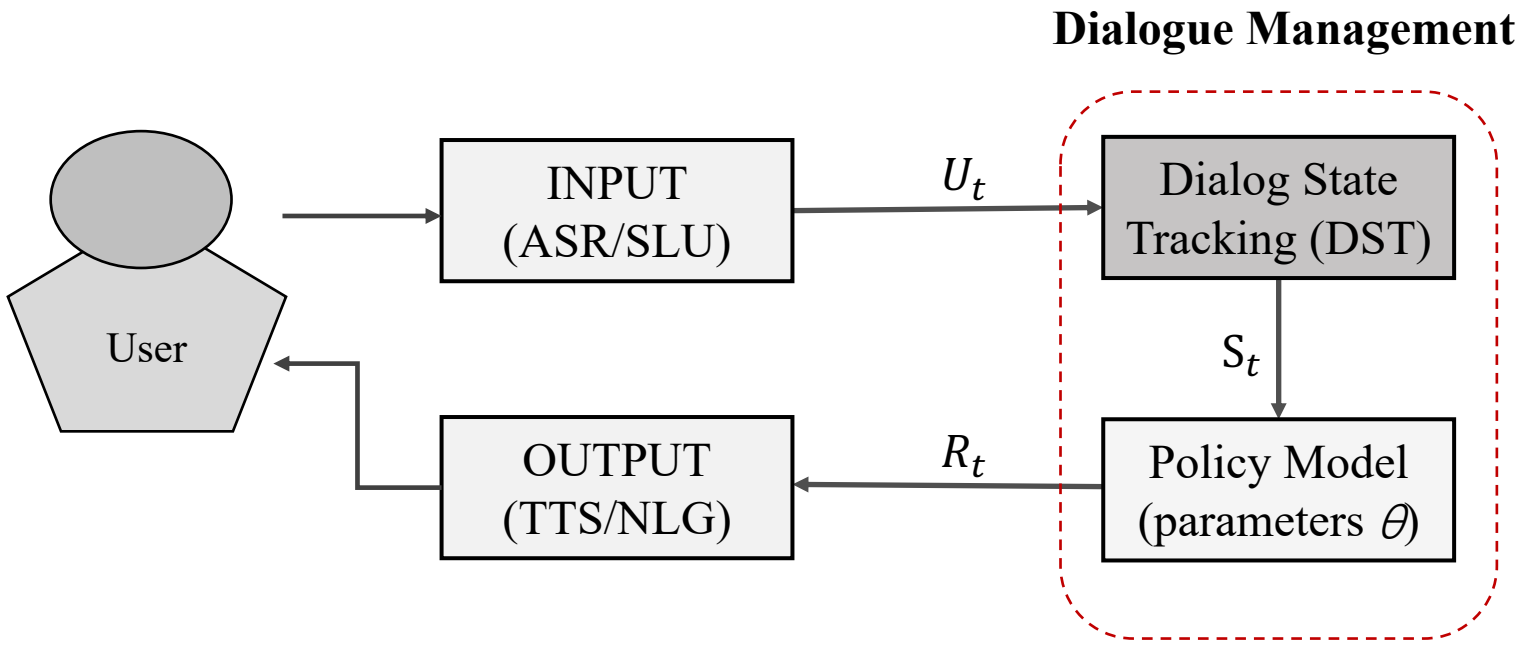
Shanghai Jiao Tong University

Outline

Dialogue System Classification

	Chat	QA	Task-oriented
context	Y	N	Y
database	N	Y	Y
annotated data	large	medium	small
popular products	Xiao Ice	IBM Watson	Amazon Echo

Task-oriented Dialogue System



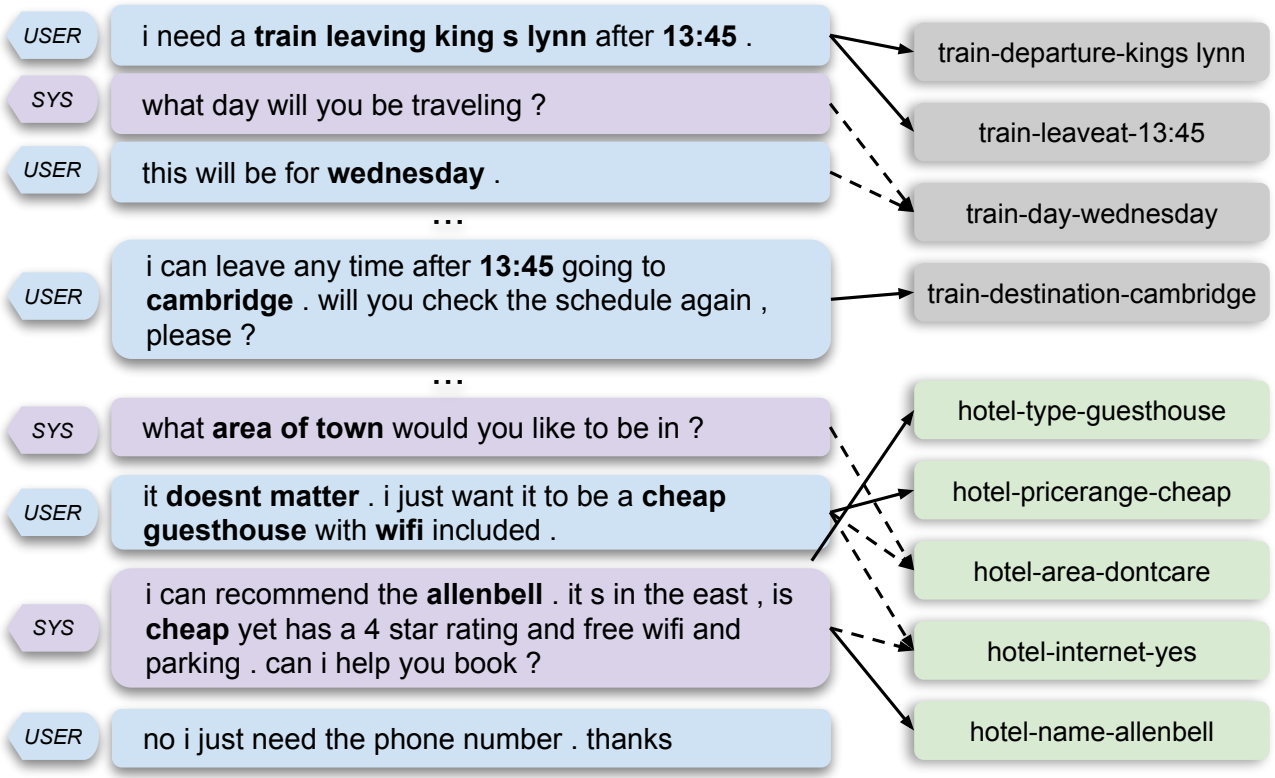
Ontology

The task is defined by the ontology.

Domains	restaurant				train		…
Slots	area	price range	food	phone	departure	…	…
Candidate Values	1. center 2. downtown …	1. cheap 2. medium 3. expensive	1. Chinese 2. Indian 3. Italian …	1. 958*** 2. 957*** …	1. cambridge 2. kings lynn …	…	…

End2End Dialogue State Tracking

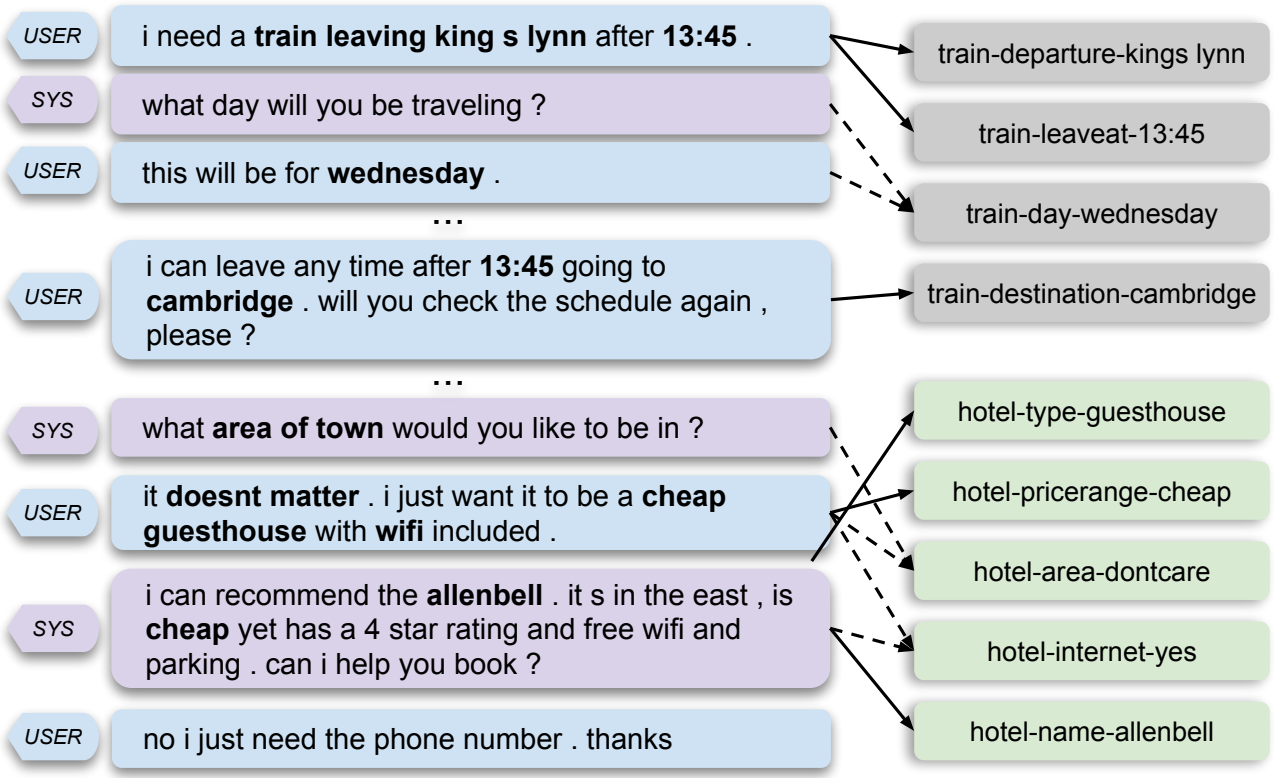
End2End DST = Dialogue-Level SLU



Input: dialogue history + ontology

Output: slot value of the ontology

Challenges



- **domain transition: from train to hotel**
- **slot confusion: many shared values between departure and destination**

Outline

1. Task Definition
2. Methodology
3. Results
4. Analysis

DST Classification

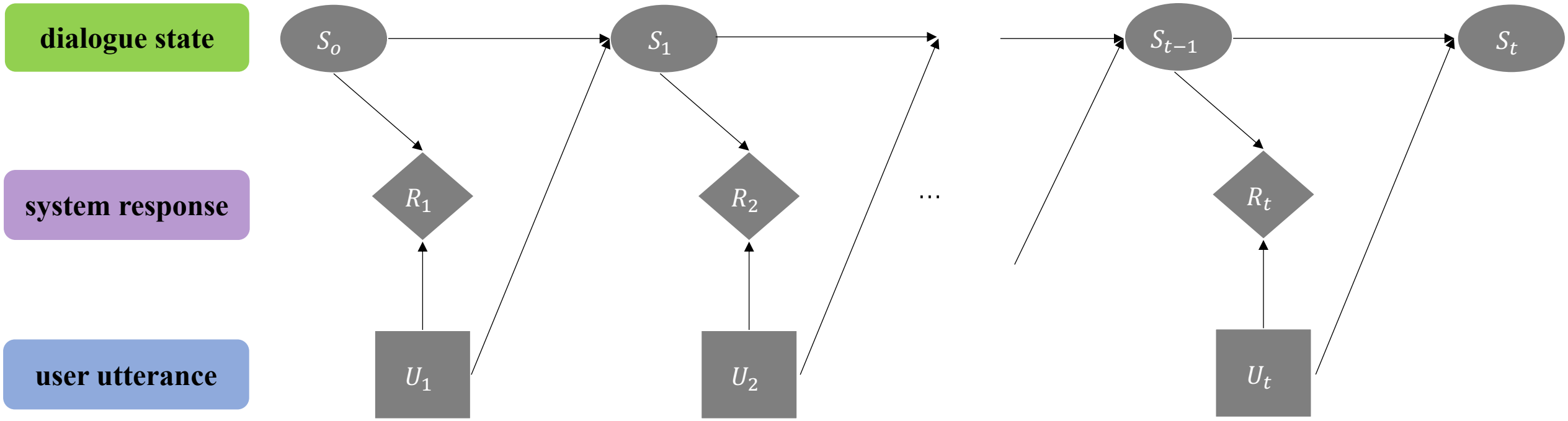
fixed-vocabulary

- all ontologies are known in advance and the model only has to pick the correct value from ontologies

open-vocabulary

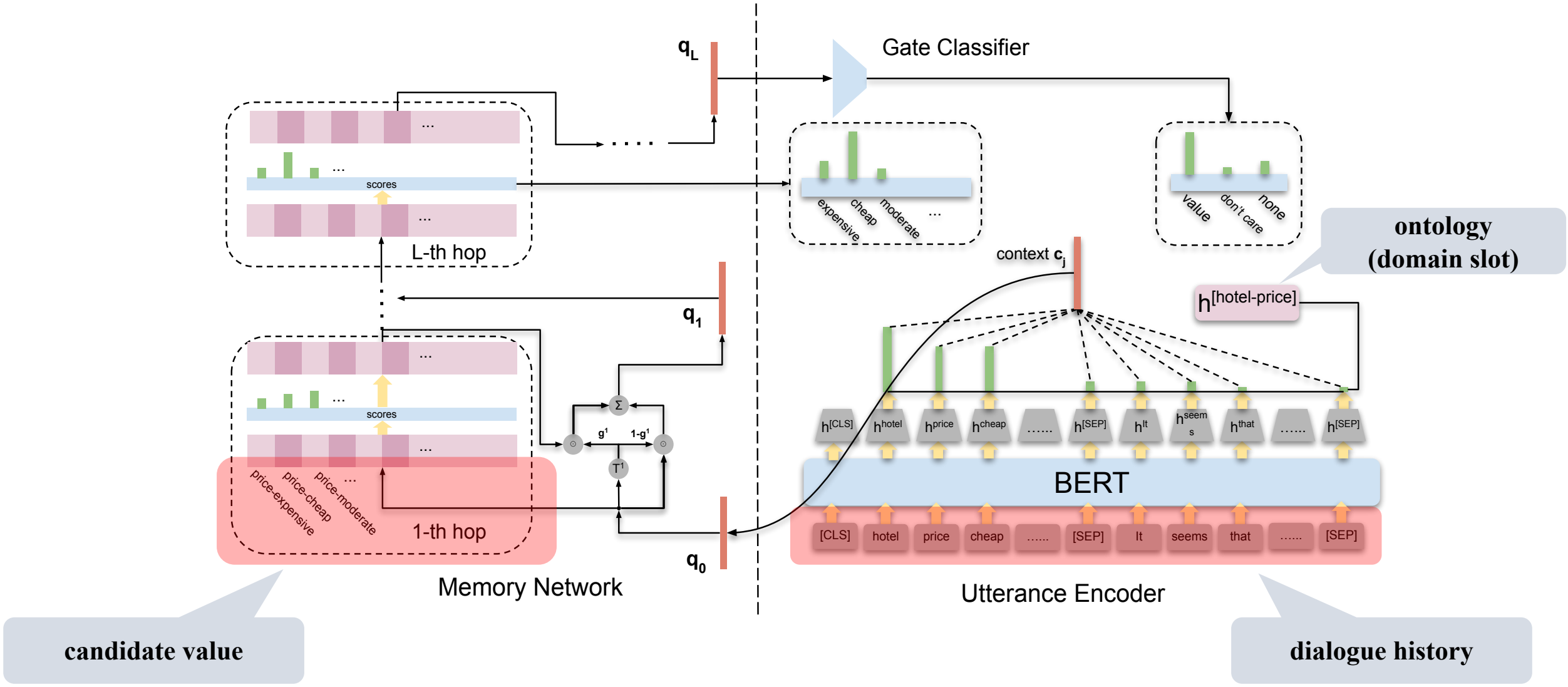
- start with zero knowledge of possible values and generate the candidate values from an open-vocabulary

Scalable Input



replace dialogue history with (S_{t-1}, R_t, U_t) in the input

Model Design



Outline

Dataset

MultiWOZ2.0 & MultiWOZ2.1

	Hotel	Train	Attraction	Restaurant	Taxi
Slots	price, type, parking, stay, day, people, area, stars, internet, name	destination, departure, day, arrive by, leave at, people	area, name, type	food, price, area, name, time, day, people	destination, departure, arrive by, leave by
Train	3381	3103	2717	3813	1654
Valid	416	484	401	438	207
Test	394	494	395	437	195

- **multi-turn dialogues: 8438**
- **average turn length: 13.68**

Metric & Setup

Metrics:

- **Joint goal accuracy: all the values of the slots are successfully predicted at each turn.**
- **Slot accuracy: the value of the corresponding slot is successfully predicted at each turn.**

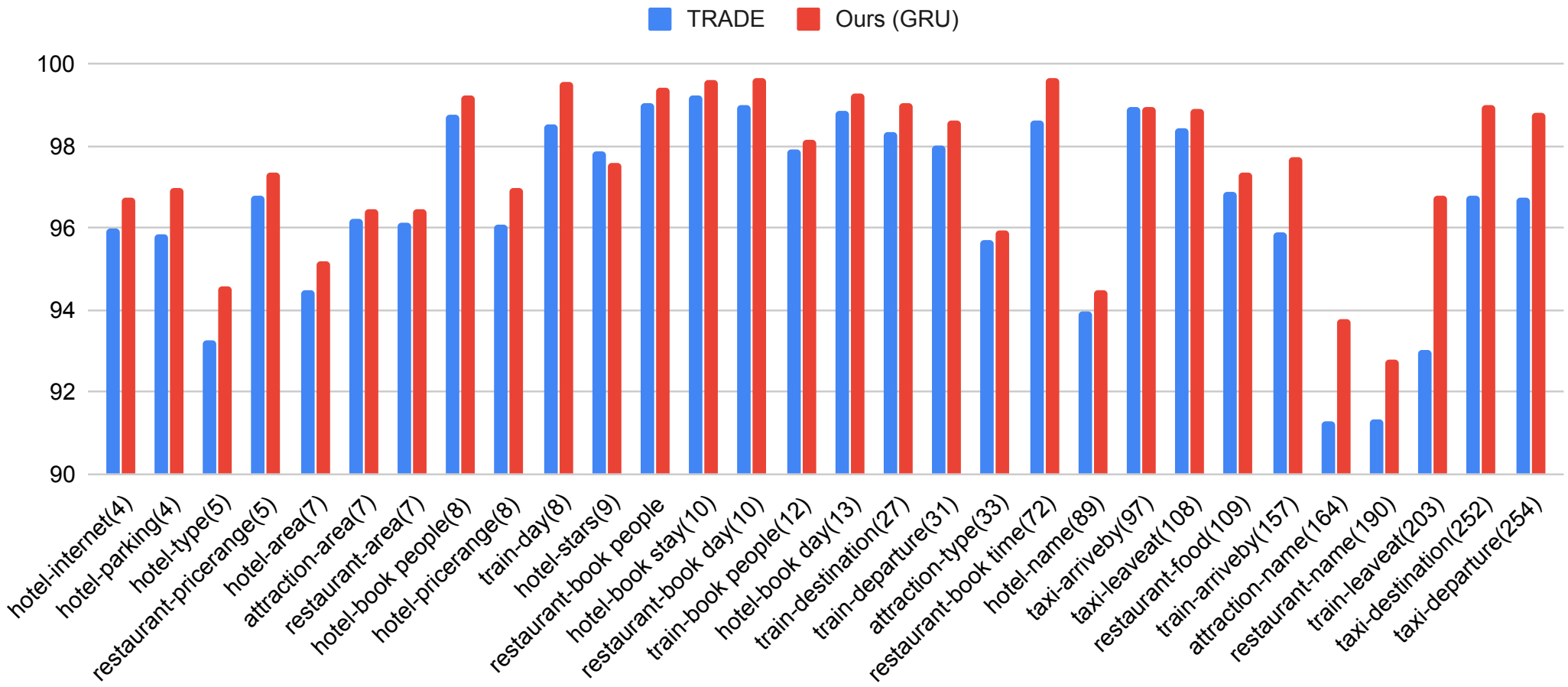
Setup:

- **Learning rate: 1e-5 for utterance encoder and 1e-4 for memory network**
- **Batch size: 16**
- **Memory hops: 2**

Joint Goal Accuracy

Models	Predefined Ontology	BERT used	MultiWoz 2.0	MultiWoz 2.1
HJST [7]	Y	N	38.40	35.55
FJST [7]	Y	N	40.20	38.00
TRADE [19]	N	N	48.60	45.60
Ours (GRU)	Y	N	**49.85**	**51.79**
SOM-DST [12]	N	Y	**51.38**	52.57
DS-DST [20]	Y	Y	-	51.21
DST-picklist [20]	Y	Y	-	53.30
SST [4]	Y	Y	51.17	55.23
Trippy [11]	N	Y	-	**55.29**
Ours (BERT)	Y	Y	50.15	52.70

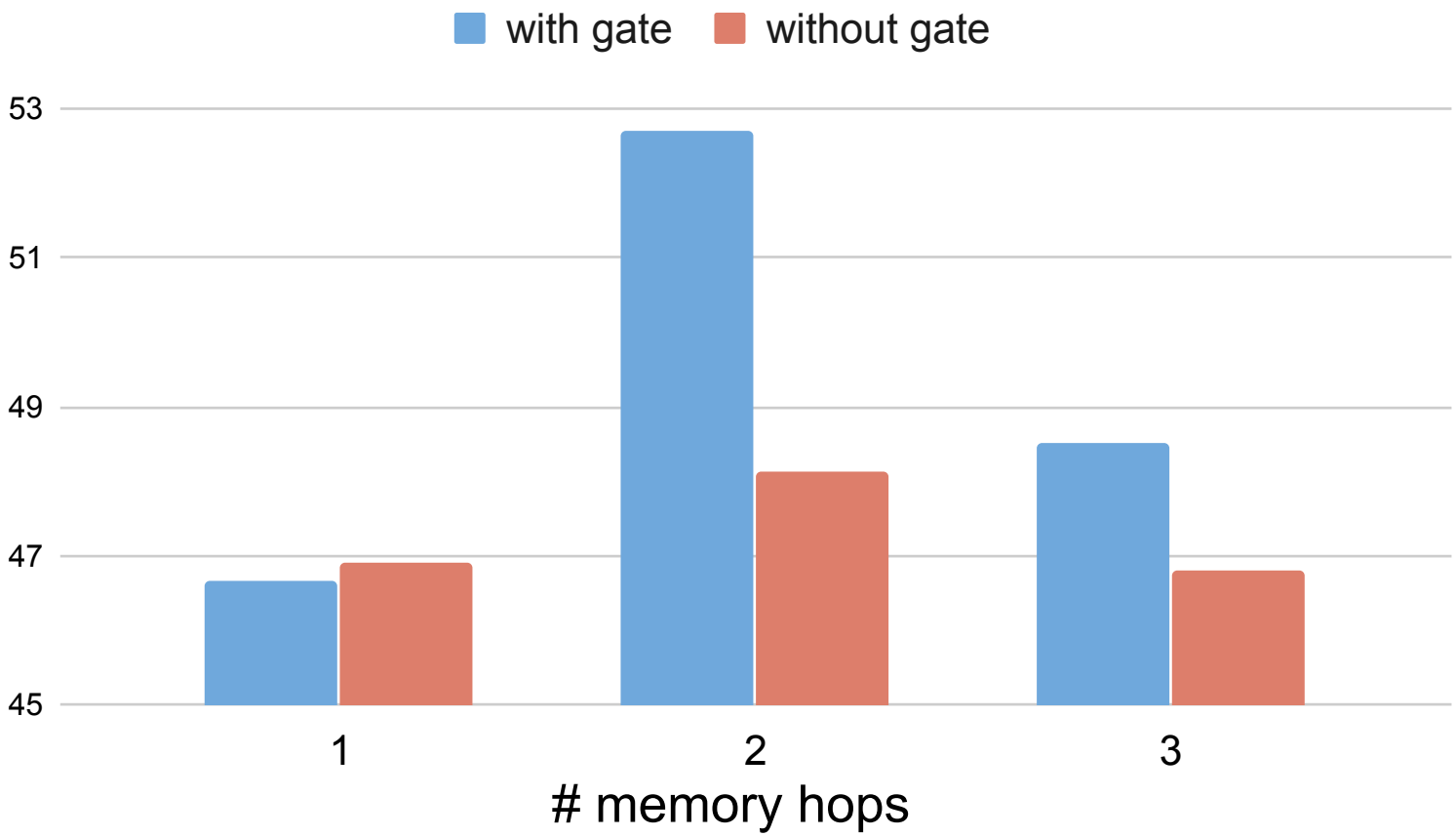
Slot Accuracy



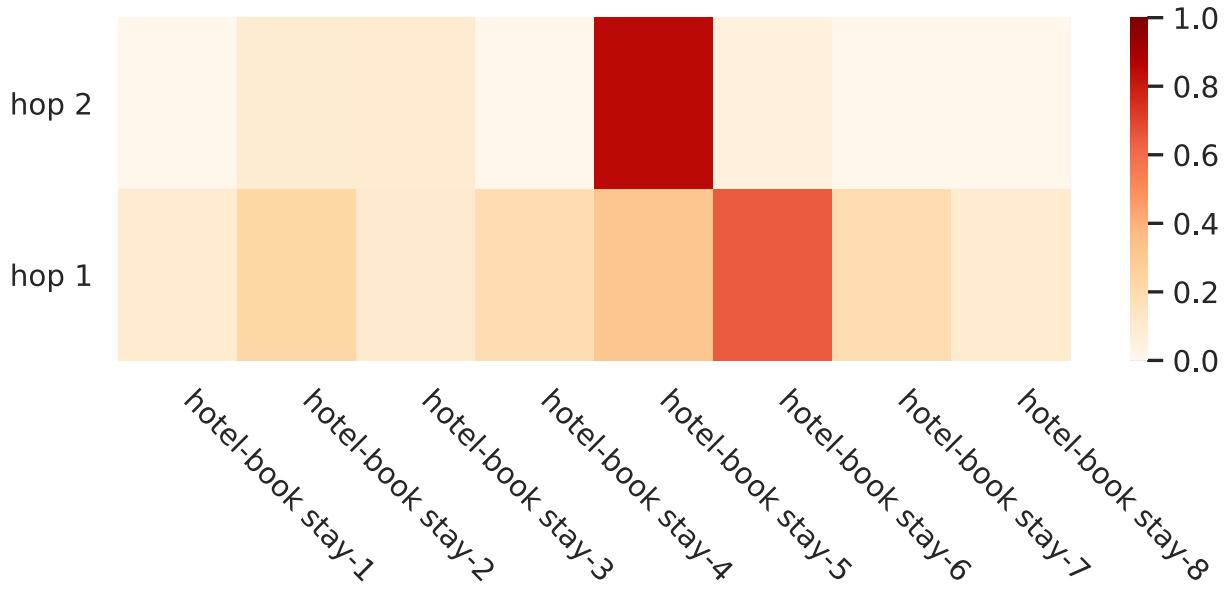
Outline

1. Task Definition
2. Methodology
3. Results
4. Analysis

Effect of Memory Gates and Memory Hops



Example Analysis



- Ontology input:
  - hotel-book stay
- Turn utterance:
  - System: i would recommend express by holiday inn cambridge . from what day should i book ?
  - User: starting saturday . i need **5 nights for 6 people** by the way .

Q&A

Thanks for your attention!