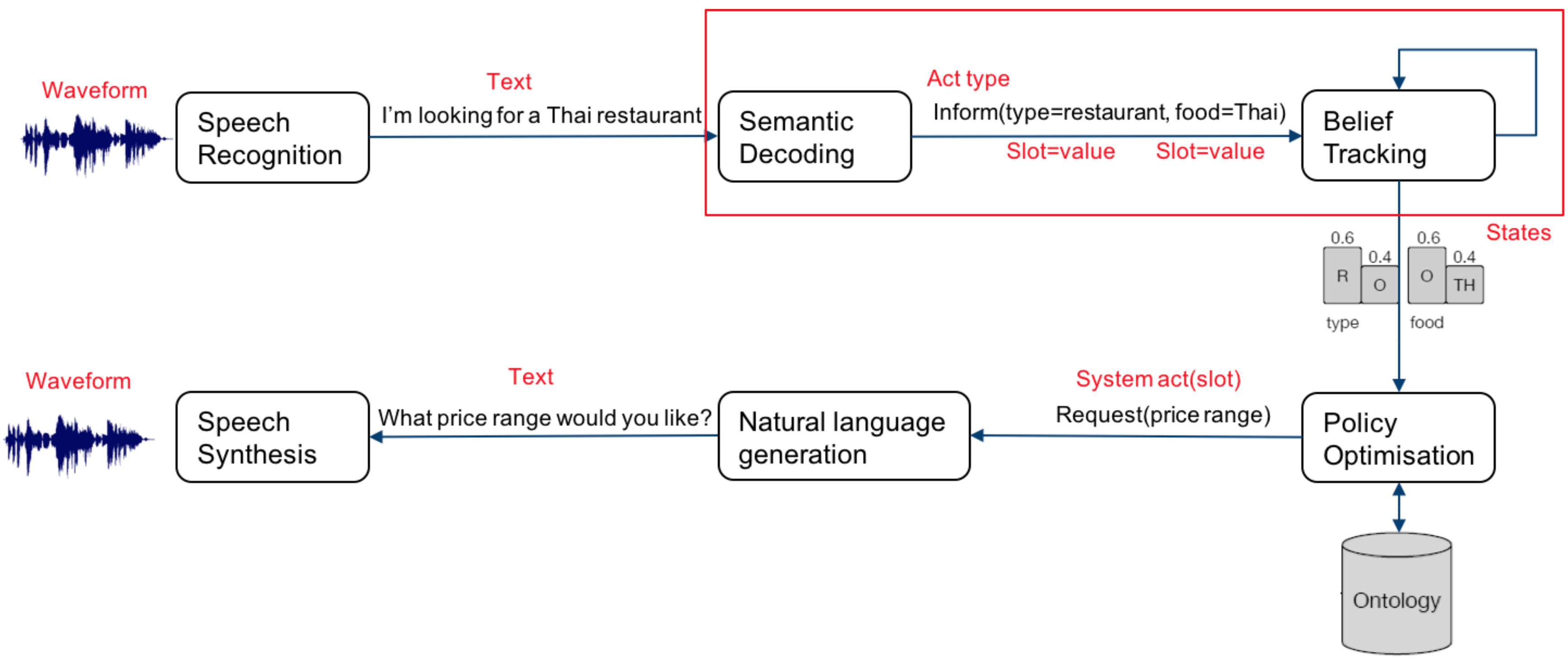
Large-Scale Multi-Domain Belief Tracking with Knowledge Sharing

Osman Ramadan, Paweł Budzianowski, Milica Gašić

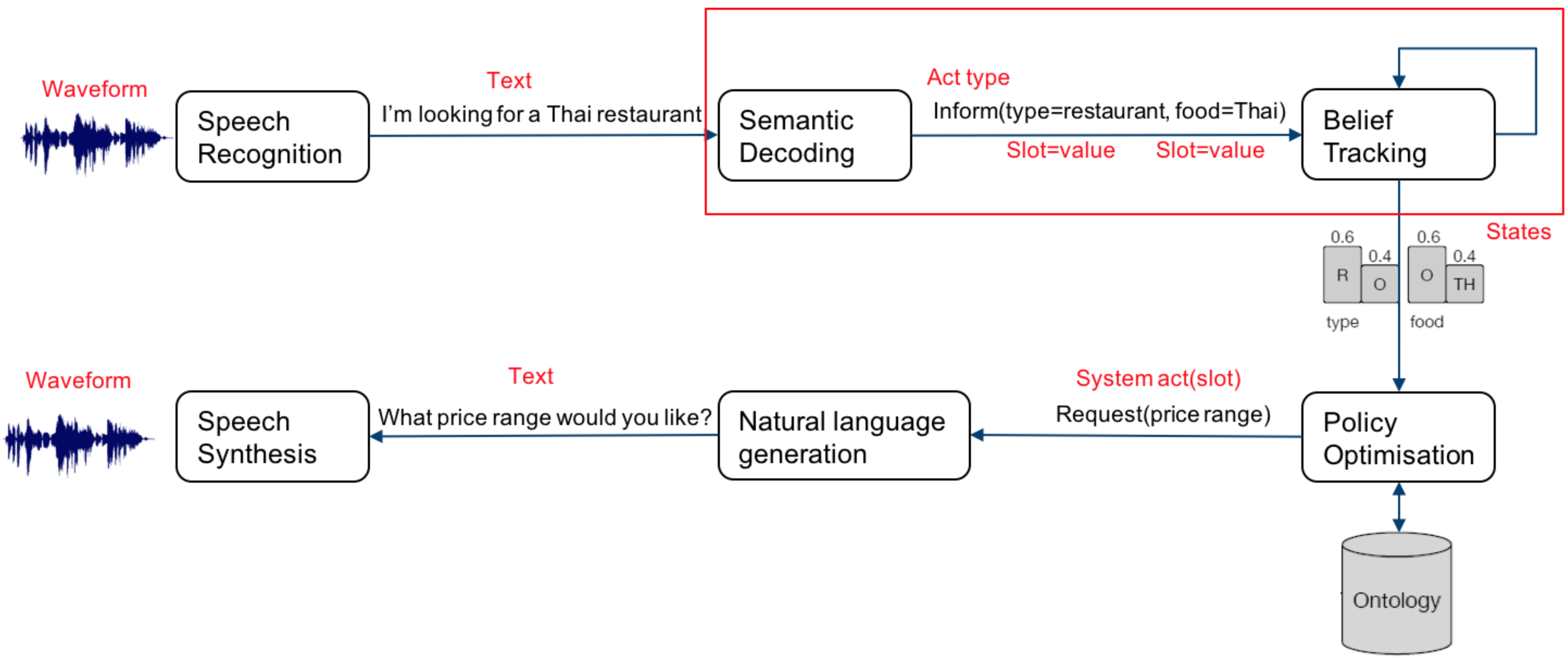
Cambridge University, Engineering Department

{oior2, pfb30, mg460}@cam.ac.uk

Main Components of Spoken Dialogue Systems

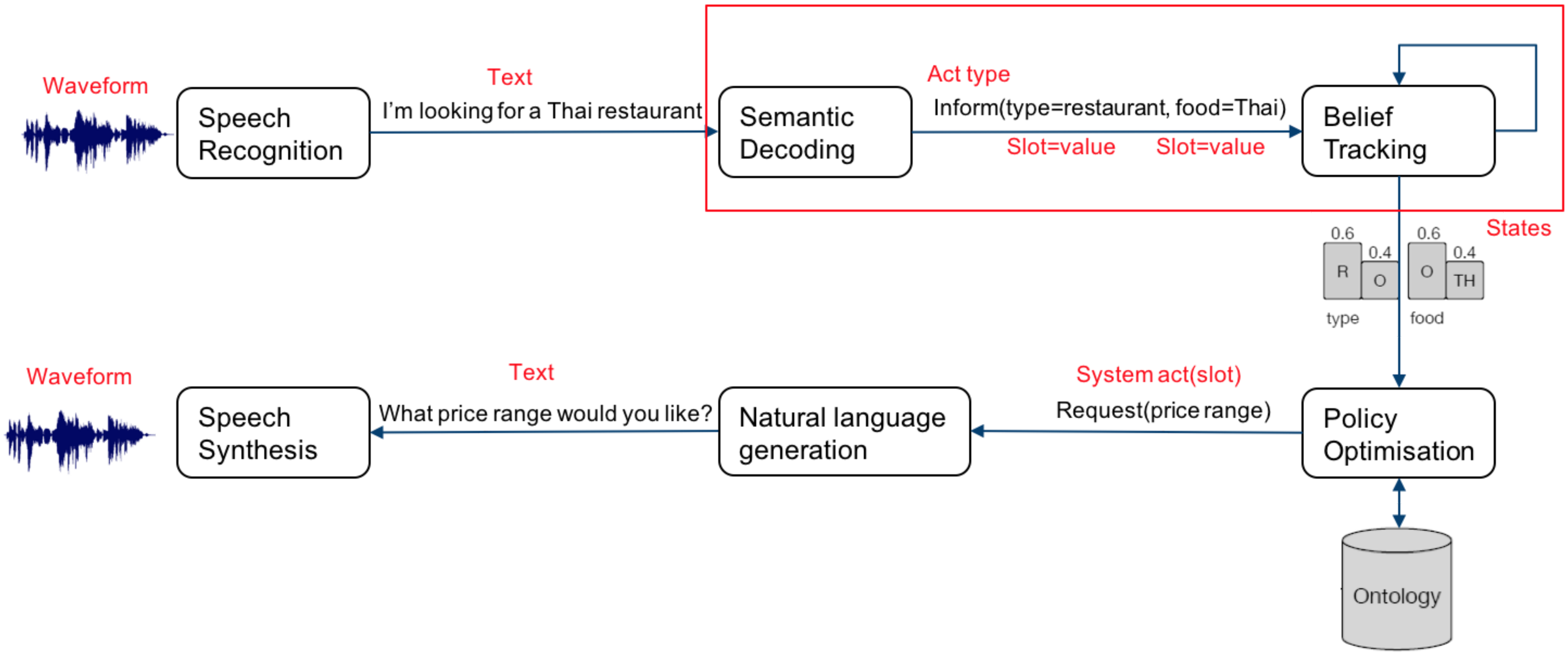


Main Components of Spoken Dialogue Systems



- Semantic decoding and belief tracking require different type of **labelled data**

Main Components of Spoken Dialogue Systems



- Semantic decoding and belief tracking require different type of **labelled data**
- Combining these two units, reduces the **amount of labelled data** required and avoid possibility of information loss in the SD stage.

Belief Tracking

Belief Tracking

`Turn 2:`
System: There are nine guesthouse hotels in various areas. What part of town are you hoping for?
User: I just need it booked for 6 people for a total of 4 nights starting from sunday.

Labels: `hotel:{internet=yes, type=guesthouse, parking=yes, pricerange=cheap Book=day, Book=people, Book=stay}`

`Turn 3:`
System: You're booked at the Alexander Bed and Breakfast, 517a coldham lane,
for 6 people for four nights starting Sunday.
User: Thank you! I'm also looking for a restaurant. Ideally an Italian place in the same price range in the centre.

Labels: `hotel:{internet=yes, type=guesthouse, parking=yes pricerange=cheap Book=day, Book=people, Book=stay}, restaurant: {area=centre, food=Italian, pricerange=cheap}`
...

Limitations of Current Belief Trackers

Limitations of Current Belief Trackers

1. The **model parameters** increase with the size of the ontology

Limitations of Current Belief Trackers

1. The **model parameters** increase with the size of the ontology
2. Many approaches rely on the **delexicalization** except for Neural Belief Tracker (NBT), Mrkšić et al 2017

Limitations of Current Belief Trackers

1. The **model parameters** increase with the size of the ontology
2. Many approaches rely on the **delexicalization** except for Neural Belief Tracker (NBT), Mrkšić et al 2017
3. Current multi-domain models do not handle **mixed domains** within a single dialogue

Limitations of Current Belief Trackers

1. The **model parameters** increase with the size of the ontology
2. Many approaches rely on the **delexicalization** except for Neural Belief Tracker (NBT), Mrkšić et al 2017
3. Current multi-domain models do not handle **mixed domains** within a single dialogue

This causes a **bottleneck in scaling** the belief tracker to larger domains and complex dialogues

Problem Formulation

Problem Formulation

1. What is in the dialogue ontology?

Problem Formulation

1. What is in the dialogue ontology?
2. What does the system output refer to?

Problem Formulation

1. What is in the dialogue ontology?
2. What does the system output refer to?
3. What does the user input refer to?

Problem Formulation

1. What is in the dialogue ontology?
2. What does the system output refer to?
3. What does the user input refer to?
4. How do we track the dialogue context?

Problem Formulation

1. What is in the dialogue ontology?
2. What does the system output refer to?
3. What does the user input refer to?
4. How do we track the dialogue context?
5. How do we handle many domains?

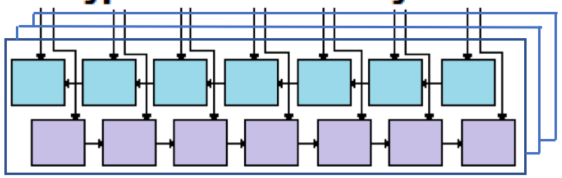
Belief Tracking with Knowledge Sharing

Belief Tracking with Knowledge Sharing

Belief Tracking with Knowledge Sharing

System

What type of food would you like



3 Bidirectional LSTMs

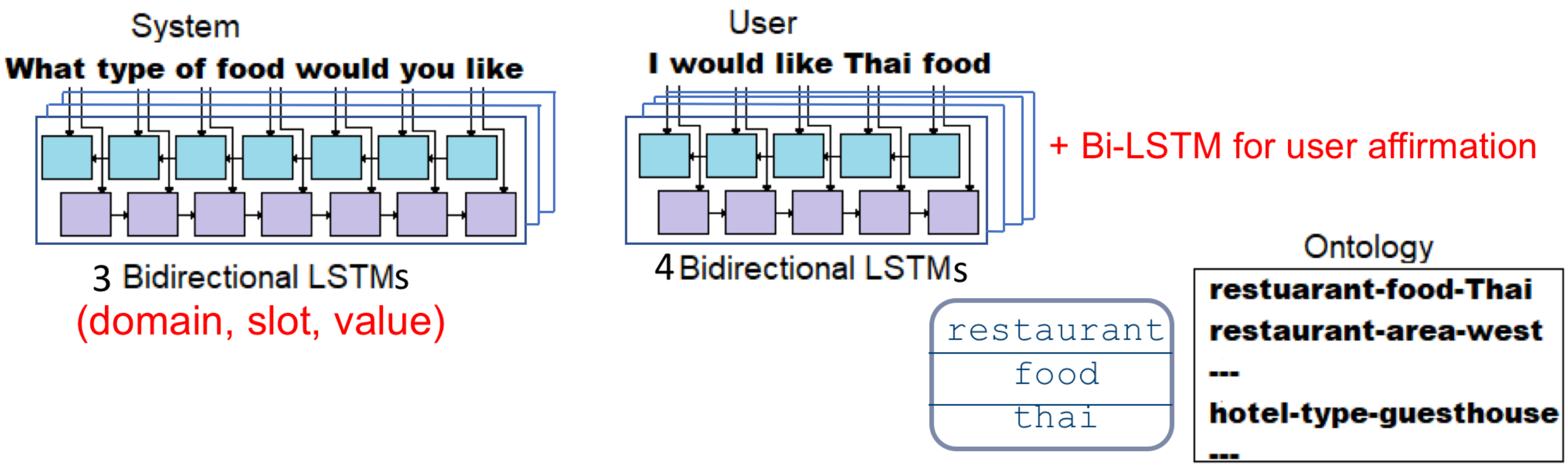
(domain, slot, value)

restaurant
food
thai

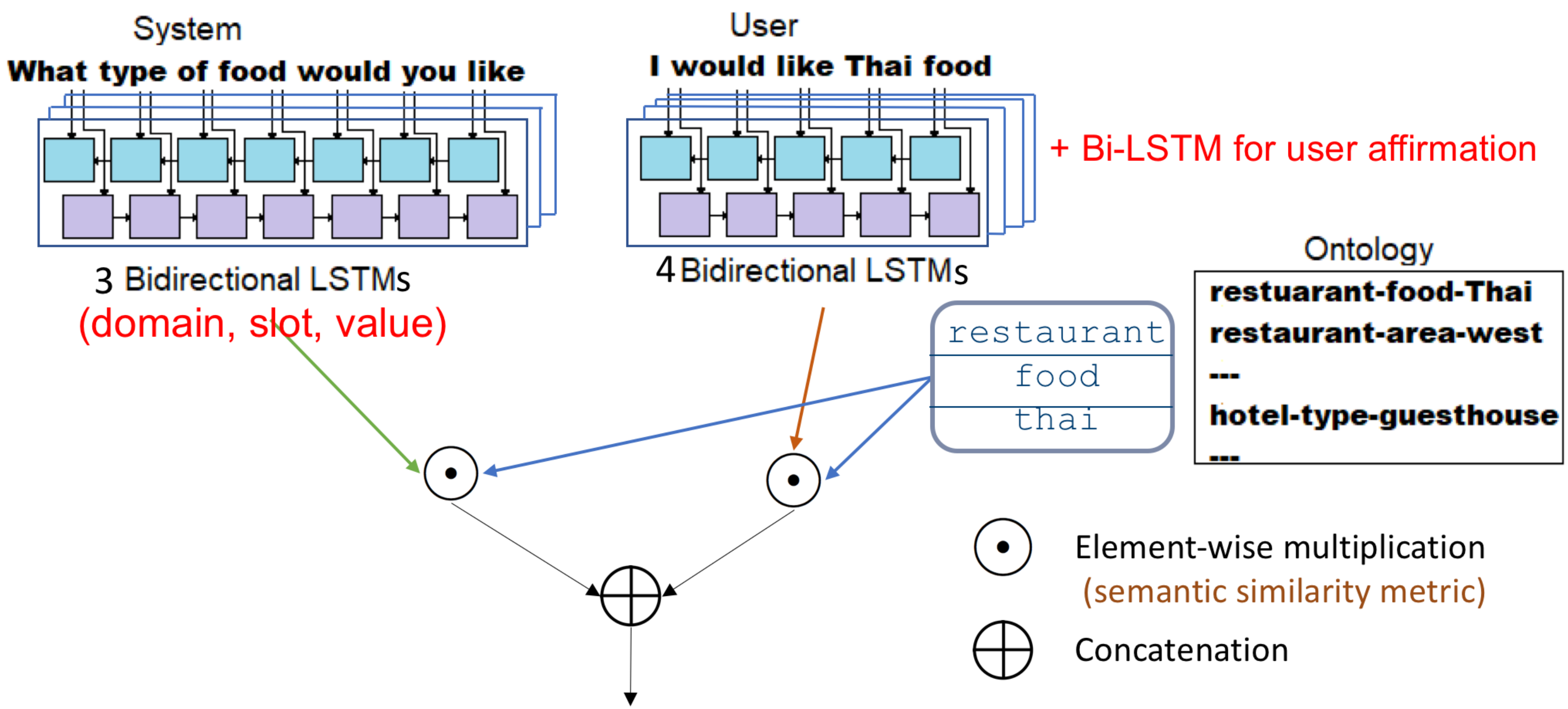
Ontology

restuarant-food-Thai
restaurant-area-west
---
hotel-type-guesthouse

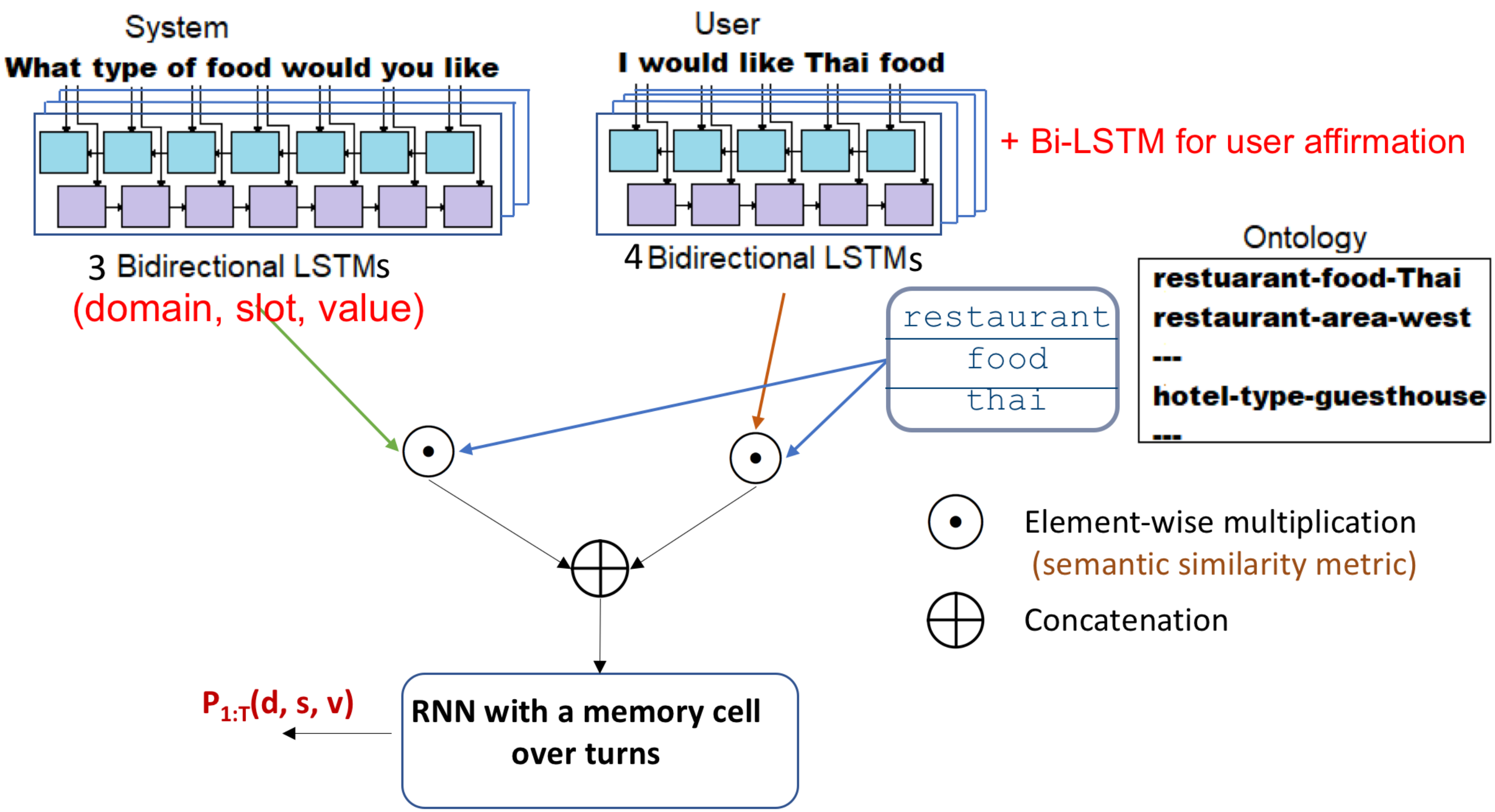
Belief Tracking with Knowledge Sharing



Belief Tracking with Knowledge Sharing



Belief Tracking with Knowledge Sharing



Belief State Update

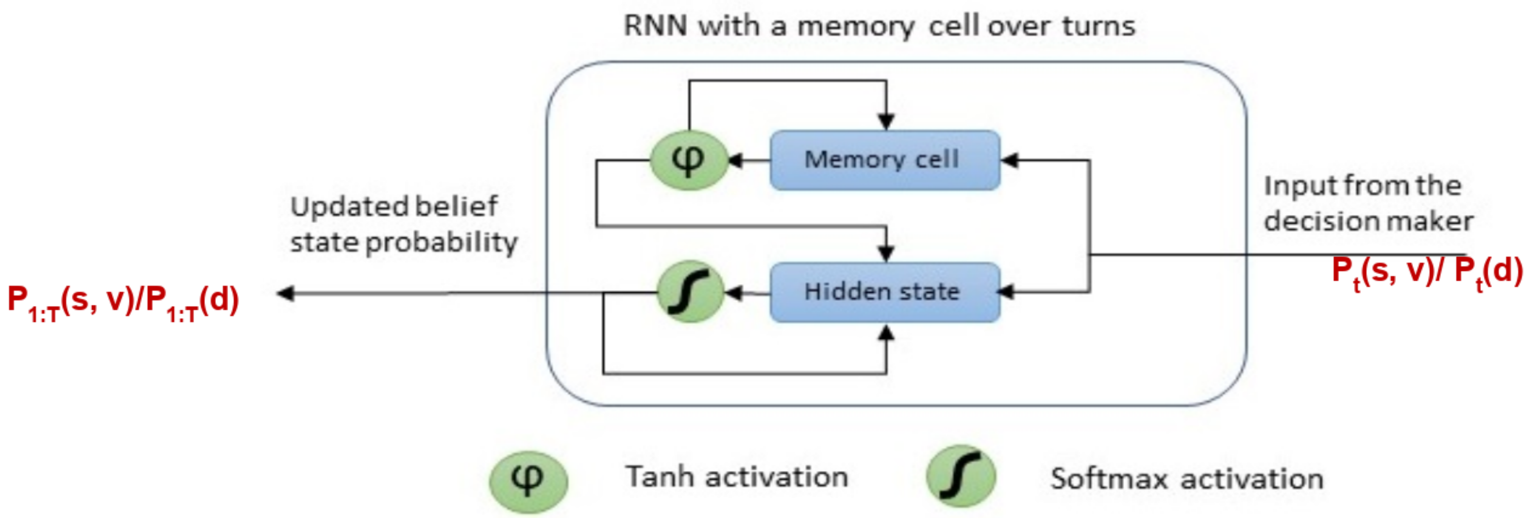
- Use a **statistical belief update** mechanism modelled by a RNN

Belief State Update

- Use a **statistical belief update** mechanism modelled by a RNN

Belief State Update

- Use a **statistical belief update** mechanism modelled by a RNN



Datasets

- Wizard of Oz framework for collecting data for belief tracking
- Amazon MTurk users given tasks to complete, access to the database
- They produce dialogues and annotate them
- Single-domain dataset WOZ 2.0 (Wen et al 2016)
- New multi-domain dataset MultiWOZ

Datasets

	WOZ 2.0	New Dataset
# of dialogues	1200	9855
# of domains	1	5
Avg. # of turns	7.45	14.30
# of slots	7	27
# of values	99	663

Results

1. Single-domain Dialogues:

	WOZ 2.0			MultiWOZ (only restaurants)		
Slot	**NBT-CNN**	**Bi-LSTM**	**CNN**	**NBT-CNN**	**Bi-LSTM**	**CNN**
Food	88.9	96.1	**96.4**	78.3	84.7	**85.3**
Price range	93.7	**98.0**	97.9	92.6	**95.6**	93.6
Area	94.3	97.8	**98.1**	78.3	82.6	**86.4**
Joint goals	84.2	85.1	**85.5**	57.7	59.9	**63.7**

Results

1. Single-domain Dialogues:

	WOZ 2.0			MultiWOZ (only restaurants)		
Slot	**NBT-CNN**	**Bi-LSTM**	**CNN**	**NBT-CNN**	**Bi-LSTM**	**CNN**
Food	88.9	96.1	**96.4**	78.3	84.7	**85.3**
Price range	93.7	**98.0**	97.9	92.6	**95.6**	93.6
Area	94.3	97.8	**98.1**	78.3	82.6	**86.4**
Joint goals	84.2	85.1	**85.5**	57.7	59.9	**63.7**

2. Multi-Domain Dialogues:

MultiWOZ (multi-domain)		
Model	**F1 score**	**Accuracy %**
Uniform Sampling	0.108	10.8
Bi-LSTM	0.876	**93.7**
CNN	**0.878**	93.2

Conclusion

1. We proposed a model with **ontology-independent** parameters

Conclusion

1. We proposed a model with **ontology-independent** parameters
2. It also achieves **state-of-the-art** results in single-domain dialogues

Conclusion

1. We proposed a model with **ontology-independent** parameters
2. It also achieves **state-of-the-art** results in single-domain dialogues
3. The model demonstrates also **great capability** in handling mixed domains dialogues

Conclusion

1. We proposed a model with **ontology-independent** parameters
2. It also achieves **state-of-the-art** results in single-domain dialogues
3. The model demonstrates also **great capability** in handling mixed domains dialogues

Future work is to test the model on **out-of-domain** tracking (zero-shot learning)

Conclusion

1. We proposed a model with **ontology-independent** parameters
2. It also achieves **state-of-the-art** results in single-domain dialogues
3. The model demonstrates also **great capability** in handling mixed domains dialogues

Future work is to test the model on **out-of-domain** tracking (zero-shot learning)

The data collection was funded through Google Faculty Award