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

Deep Learning Summer school, 05 Aug 2016
Tsung-Hsien (Shawn) Wen
Outline

- Intro
- Neural Dialogue System
- Wizard-of-Oz Data Collection
- Experiments
- Conclusion
Outline

- **Intro**
- Neural Dialogue System
- Wizard-of-Oz Data Collection
- Experiments
- Conclusion
The two paradigms

- Goal-oriented Dialogue Systems
The two paradigms

- Goal-oriented Dialogue Systems
  - Help the user to accomplish domain tasks
The two paradigms

- Goal-oriented Dialogue Systems
  - Help the user to accomplish domain tasks
  - Domain specific, hard to collect data
The two paradigms

- **Goal-oriented Dialogue Systems**
  - Help the user to accomplish domain tasks
  - Domain specific, hard to collect data

- **Current Systems**
  - Modular, highly handcrafted, restricted ability
The two paradigms

- Goal-oriented Dialogue Systems
  - Help the user to accomplish domain tasks
  - Domain specific, hard to collect data

- Current Systems
  - Modular, highly handcrafted, restricted ability

Can we train a dialogue system on a small dataset w/ a minimal amount of handcrafting?
The two paradigms

- Goal-oriented Dialogue Systems
  - Help the user to accomplish domain tasks
  - Domain specific, hard to collect data

- Current Systems
  - Modular, highly handcrafted, restricted ability

- Can we train a dialogue system on a small dataset w/ a minimal amount of handcrafting?
- How can we collect data to train this model?
Outline

- Intro
- Neural Dialogue System
- Wizard-of-Oz Data Collection
- Experiments
- Conclusion
Traditional Dialogue Systems

Dialogue System

Language Understanding

Language Generation

Dialogue Manager

Ontology

DB
Neural Dialogue Systems
Can I have food?

Korean 0.7
British 0.2
French 0.1

Belief Tracker

MySQL query: “Select * where food=Korean”

Policy Network

Generation Network

Copy field

Database Operator

Intent Network

Can I have <v.food>
Can I have \texttt{<v.food>}?

\texttt{<v.name>} serves great \texttt{<v.food>}.
Can I have <v.food>

Korean 0.7
British 0.2
French 0.1
...

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

Language Grounding
Can I have <v.food>

Korean 0.7
British 0.2
French 0.1

MySQL query: "Select * where food=Korean"

Intent Network

Generation Network

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

Belief Tracker

Database Accessing

Database Operator
Can I have <v.food>

Belief Tracker

Korean 0.7
British 0.2
French 0.1
...

Policy Network

MySQL query:
"Select * where food=Korean"

Decision Making

Belief Tracker

MySQL query:
"Select * where food=Korean"

Belief Tracker

Database Operator

Intent Network

Can I have <v.food>

Generation Network

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

Generation Network

Belief Tracker

Database

Seven days
Curry Prince
Nirala
Royal Standard
Little Seoul

DB pointer

x

p_t

q_t

Decision Making

Database Operator

Intent Network

Can I have <v.food>

Generation Network

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

Belief Tracker
Can I have <v.food>

Korean 0.7
British 0.2
French 0.1
...

MySQL query: "Select * where food=Korean"

Policy Network

Copy field

DB pointer

Database Operator

Intent Network

Generation Network

Copy field

Belief Tracker

Database

Seven days
Curry Prince
Nirala
Royal Standard
Little Seoul

Can I have <v.food>
Outline

- Intro
- Neural Dialogue System
- **Wizard-of-Oz Data Collection**
- Experiments
- Conclusion
Wizard of Oz Data Collection

- Online parallel version of WOZ on MTurk
Wizard of Oz Data Collection

- Online parallel version of WOZ on MTurk
  - Randomly hire a worker to be user/wizard.
Wizard of Oz Data Collection

- Online parallel version of WOZ on MTurk
  - Randomly hire a worker to be user/wizard.
  - Task: Enter an appropriate response for ONE TURN.
Wizard of Oz Data Collection

- Online parallel version of WOZ on MTurk
  - Randomly hire a worker to be user/wizard.
  - Task: Enter an appropriate response for ONE TURN.
  - Repeat the process until all dialogues are finished.
Wizard of Oz Data Collection

- Online parallel version of WOZ on MTurk
  - Randomly hire a worker to be user/wizard.
  - Task: Enter an appropriate response for ONE TURN.
  - Repeat the process until all dialogues are finished.

- Advantage:
Wizard of Oz Data Collection

- Online parallel version of WOZ on MTurk
  - Randomly hire a worker to be user/wizard.
  - Task: Enter an appropriate response for ONE TURN.
  - Repeat the process until all dialogues are finished.

- Advantage:
  - Avoid latency, parallelisable, cheap
Task:
Find a restaurant, cheap, Korean, North
Ask phone number
Hi, I want a cheap Korean restaurant.
Task:
Find a restaurant, cheap, Korean, North
Ask phone number

Hi, I want a cheap Korean restaurant.
Hi, I want a **cheap Korean restaurant**.
Hi, I want a cheap Korean restaurant.

What user wants?

<table>
<thead>
<tr>
<th>Food</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>Cheap</td>
</tr>
<tr>
<td>Area</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Search Table

<table>
<thead>
<tr>
<th>Little Seoul</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Korea</td>
<td>...</td>
</tr>
</tbody>
</table>

Task:
Find a restaurant, cheap, Korean, North
Ask phone number
Hi, I want a cheap Korean restaurant.

What area are you looking for?

What user wants?

<table>
<thead>
<tr>
<th>Food</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>Cheap</td>
</tr>
<tr>
<td>Area</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Search Table

<table>
<thead>
<tr>
<th>Little Seoul</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Korea</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>
Hi, I want a **cheap Korean restaurant**.

**What area are you looking for?**
Wizard of Oz Data Collection

Task: Find a restaurant, cheap, Korean, North
Ask phone number

Hi, I want a cheap Korean restaurant.

What area are you looking for?

Somewhere in the north.
Hi, I want a **cheap Korean restaurant**.

What **area** are you looking for?

**Somewhere in the north.**

---

**Task:**
- Find a restaurant, cheap, Korean, North
- Ask phone number

**What user wants?**

<table>
<thead>
<tr>
<th>Food</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>Cheap</td>
</tr>
<tr>
<td>Area</td>
<td>North</td>
</tr>
</tbody>
</table>
Hi, I want a cheap Korean restaurant.

What area are you looking for?

Somewhere in the north.
Hi, I want a **cheap Korean restaurant**.

**What area** are you looking for?

*Somewhere in the north.*

**Little Seoul** is nice one in the north.

<table>
<thead>
<tr>
<th>What user wants?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Food</strong></td>
<td>Korean</td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td>Cheap</td>
</tr>
<tr>
<td><strong>Area</strong></td>
<td>North</td>
</tr>
</tbody>
</table>

**Search Table**

<table>
<thead>
<tr>
<th>Little Seoul</th>
<th></th>
</tr>
</thead>
</table>
Hi, I want a cheap Korean restaurant.

What area are you looking for?

Somewhere in the north.

Little Seoul is nice one in the north.

What is the phone number?

Its phone number is 01223456789.

Thank you very much, good bye.

Thank you for using the system.
Hi, I want a cheap Korean restaurant.

What area are you looking for?

Somewhere in the north.

Little Seoul is nice one in the north.

What is the phone number?

Its phone number is 01223456789.

Thank you very much, good bye.

Thank you for using the system.
Data Statistics

- Ontology:
Data Statistics

- Ontology:
  - Cambridge restaurant domain, 99 venues.
Data Statistics

- **Ontology:**
  - Cambridge restaurant domain, 99 venues.
  - 3 informable slots: area, price range, food type
  - 3 requestable slots: address, phone, postcode
Data Statistics

- Ontology:
  - Cambridge restaurant domain, 99 venues.
  - 3 informable slots: area, price range, food type
  - 3 requestable slots: address, phone, postcode

- Dataset
Data Statistics

- Ontology:
  - Cambridge restaurant domain, 99 venues.
  - 3 informable slots: area, price range, food type
  - 3 requestable slots: address, phone, postcode

- Dataset
  - 676 dialogues, ~2750 turns
Data Statistics

- **Ontology:**
  - Cambridge restaurant domain, 99 venues.
  - 3 informable slots: area, price range, food type
  - 3 requestable slots: address, phone, postcode

- **Dataset**
  - 676 dialogues, ~2750 turns
  - 3000 HITS, takes 3 days, costs ~400 USD
Data Statistics

- **Ontology:**
  - Cambridge restaurant domain, 99 venues.
  - 3 informable slots: area, price range, food type
  - 3 requestable slots: address, phone, postcode

- **Dataset**
  - 676 dialogues, ~2750 turns
  - 3000 HITS, takes 3 days, costs ~400 USD
  - Data cleaning takes 2-3 days for one person
Outline

- Intro
- Neural Dialogue System
- Wizard-of-Oz Data Collection
- Experiments
- Conclusion
Experiments

- Experimental details
  - Train/valid/test: 3/1/1
  - SGD, l2 regularisation, early stopping, gradient clip=1
  - Hidden size = 50, Vocab size: ~500
Experiments

- Experimental details
  - Train/valid/test: 3/1/1
  - SGD, l2 regularisation, early stopping, gradient clip=1
  - Hidden size = 50, Vocab size: ~500

- Two stage training:
  - Training trackers with label cross entropy
  - Training other parts with response cross entropy
Experiments

- Experimental details
  - Train/valid/test: 3/1/1
  - SGD, l2 regularisation, early stopping, gradient clip=1
  - Hidden size = 50, Vocab size: ~500

- Two stage training:
  - Training trackers with label cross entropy
  - Training other parts with response cross entropy

- Decoding
  - Beam search w/ beam width 10
  - Decode with average word likelihood
## Human evaluation

### System Comparison

<table>
<thead>
<tr>
<th>Metric</th>
<th>N2N</th>
<th>Modular</th>
<th>Tie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj. Success</td>
<td>96.95%</td>
<td>95.12%</td>
<td>-</td>
</tr>
<tr>
<td>Avg. # of Turn</td>
<td>3.95</td>
<td>4.54</td>
<td>-</td>
</tr>
</tbody>
</table>

**Comparisons(%)**

<table>
<thead>
<tr>
<th>Metric</th>
<th>N2N</th>
<th>Modular</th>
<th>Tie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naturalness</td>
<td>46.95*</td>
<td>25.61</td>
<td>27.44</td>
</tr>
<tr>
<td>Comprehension</td>
<td>45.12*</td>
<td>21.95</td>
<td>32.93</td>
</tr>
<tr>
<td>Preference</td>
<td>50.00*</td>
<td>24.39</td>
<td>25.61</td>
</tr>
<tr>
<td>Performance</td>
<td>43.90*</td>
<td>25.61</td>
<td>30.49</td>
</tr>
</tbody>
</table>

* $p < 0.005$,  
# of comparisons: 164
Example dialogues

**Human** : Restaurant in any area that serves Vietnamese food
**Machine** : thanh binh serves vietnamese food type, is there anything else i can help you with?
**Human** : What is the phone number and post code
**Machine** : the number is 01223 362456 and the area code is c.b 3, 0 a.f. is there anything else i can do for you?
**Human** : Good bye
**Machine** : thank you for using the cambridge restaurant system.

**Human** : Hi, I’m hungry for some Indonesian food
**Machine** : i am sorry , there are no indonesian restaurants . would you like a different food?
**Human** : Oh no. How about Indian?
**Machine** : there are 9 restaurants serving indian types of food . do you have a preference on side of town?
**Human** : East side please
**Machine** : sitar tandoori is a indian restaurant in the east side of town . would you like their information?
**Human** : Yes, the address and phone number please
**Machine** : sitar tandoori is located at 43 high street cherry hinton cherry hinton and their phone number is 01223 249955.
**Human** : Okay, thank you
**Machine** : thank you , goodbye.
Example dialogues

<table>
<thead>
<tr>
<th>Human</th>
<th>Machine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restaurant in any area that serves Vietnamese food.</td>
<td>thanh binh serves vietnamese food type. is there anything else i can help you with?</td>
</tr>
<tr>
<td>What is the phone number and post code.</td>
<td>the number is 01223 362456 and the area code is c.b 3, 0 a.f. is there anything else i can do for you?</td>
</tr>
<tr>
<td>Good bye.</td>
<td>thank you for using the cambridge restaurant system.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human</th>
<th>Machine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hi, I’m hungry for some Indonesian food.</td>
<td>i am sorry, there are no indonesian restaurants would you like a different food?</td>
</tr>
<tr>
<td>Oh no. How about Indian?</td>
<td>there are 9 restaurants serving indian types of food. do you have a preference on side of town?</td>
</tr>
<tr>
<td>East side please.</td>
<td>sitar tandoori is a indian restaurant in the east side of town would you like their information?</td>
</tr>
<tr>
<td>Yes, the address and phone number please.</td>
<td>sitar tandoori is located at 43 high street cherry hinton cherry hinton and their phone number is 01223 249955.</td>
</tr>
<tr>
<td>Okay, thank you.</td>
<td>thank you, goodbye.</td>
</tr>
</tbody>
</table>
### Example dialogues

<table>
<thead>
<tr>
<th>Human</th>
<th>Restaurant in any area that serves Vietnamese food</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine</td>
<td>thanh binh serves vietnamese food type , is there anything else i can help you with ?</td>
</tr>
<tr>
<td>Human</td>
<td>What is the phone number and post code</td>
</tr>
<tr>
<td>Machine</td>
<td>the number is 01223 362456 and the area code is c.b 3, 0 a.f . is there anything else i can do for you ?</td>
</tr>
<tr>
<td>Human</td>
<td>Good bye</td>
</tr>
<tr>
<td>Machine</td>
<td>thank you for using the cambridge restaurant system .</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human</th>
<th>Hi, I’m hungry for some Indonesian food</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine</td>
<td>i am sorry , there are no indonesian restaurants . would you like a different food ?</td>
</tr>
<tr>
<td>Human</td>
<td>Oh no. How about Indian?</td>
</tr>
<tr>
<td>Machine</td>
<td>there are 9 restaurants serving indian types of food . do you have a preference on side of town ?</td>
</tr>
<tr>
<td>Human</td>
<td>East side please</td>
</tr>
<tr>
<td>Machine</td>
<td>sitar tandoori is a indian restaurant in the east side of town . would you like their information ?</td>
</tr>
<tr>
<td>Human</td>
<td>Yes, the address and phone number please</td>
</tr>
<tr>
<td>Machine</td>
<td>sitar tandoori is located at 43 high street cherry hinton cherry hinton and their phone number is 01223 249955 .</td>
</tr>
<tr>
<td>Human</td>
<td>Okay, thank you</td>
</tr>
<tr>
<td>Machine</td>
<td>thank you , goodbye .</td>
</tr>
</tbody>
</table>
Visualising action embedding
Outline

- Intro
- Neural Dialogue System
- Wizard-of-Oz Data Collection
- Experiments
- Conclusion
Conclusion

- An end-to-end trainable task-oriented dialogue system architecture is proposed.
Conclusion

- An end-to-end trainable task-oriented dialogue system architecture is proposed.
- A complementary WOZ data collection is also proposed (no latency, parallel, cheap).
Conclusion

- An end-to-end trainable task-oriented dialogue system architecture is proposed.
- A complementary WOZ data collection is also proposed (no latency, parallel, cheap).
- Results show that it can learn from human-human conversations and help users to complete tasks.
The paper

Thank you! Questions?

Tsung-Hsien Wen is supported by a studentship funded by Toshiba Research Europe Ltd, Cambridge Research Laboratory
## Response Generation Task

<table>
<thead>
<tr>
<th>Model</th>
<th>Match (%)</th>
<th>Success (%)</th>
<th>BLEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq2Seq [Sutskever et al, 2014]</td>
<td>-</td>
<td>-</td>
<td>0.1718</td>
</tr>
<tr>
<td>HRED [Serban et al, 2015]</td>
<td>-</td>
<td>-</td>
<td>0.1861</td>
</tr>
<tr>
<td>Our model w/o req. trackers</td>
<td>89.70</td>
<td>30.60</td>
<td>0.1799</td>
</tr>
<tr>
<td>Our full model</td>
<td>86.34</td>
<td>75.16</td>
<td>0.2313</td>
</tr>
<tr>
<td>Our full model + attention</td>
<td>90.88</td>
<td>80.02</td>
<td>0.2388</td>
</tr>
</tbody>
</table>
Example dialogues

<table>
<thead>
<tr>
<th>Human</th>
<th>hello</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine</td>
<td>what cuisine would you like?</td>
</tr>
<tr>
<td>Human</td>
<td>afghan</td>
</tr>
<tr>
<td>Machine</td>
<td>i am sorry, there are no afghan restaurants. would you like something else?</td>
</tr>
<tr>
<td>Human</td>
<td>korean</td>
</tr>
<tr>
<td>Machine</td>
<td>little seoul serves korean food, it is located at 108 regent street city centre.</td>
</tr>
<tr>
<td>Human</td>
<td>is it expensive?</td>
</tr>
<tr>
<td>Machine</td>
<td>little seoul serves korean food and is in the expensive price range.</td>
</tr>
<tr>
<td>Human</td>
<td>what area is that</td>
</tr>
<tr>
<td>Machine</td>
<td>little seoul is located in the centre area of town. would you like their number?</td>
</tr>
<tr>
<td>Human</td>
<td>yes</td>
</tr>
<tr>
<td>Machine</td>
<td>the telephone number is 01223 308681. is there anything else i can help you with today?</td>
</tr>
<tr>
<td>Human</td>
<td>no goodbye</td>
</tr>
<tr>
<td>Machine</td>
<td>thank you for using the cambridge restaurant system, goodbye.</td>
</tr>
</tbody>
</table>
Example dialogues

| Human       | hello                           |
| Machine    | what cuisine would you like?    |
| Human      | afghan                          |
| Machine    | i am sorry, there are no afghan restaurants. would you like something else? |
| Human      | korean                          |
| Machine    | little seoul serves korean food, it is located at 108 regent street city centre. |
| Human      | is it expensive?                |
| Machine    | little seoul serves korean food and is in the expensive price range. |
| Human      | what area is that               |
| Machine    | little seoul is located in the centre area of town. would you like their number? |
| Human      | yes                             |
| Machine    | the telephone number is 01223 308681. is there anything else i can help you with today? |
| Human      | no goodbye                      |
| Machine    | thank you for using the cambridge restaurant system, goodbye. |
Jordan RNN-CNN belief trackers

Jordan RNN

Output layer

Hidden layer

User turn $t$

System turn $t-1$

Turn $t$

Input layer

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation

User turn $t$

System turn $t-1$

Input layer

Output layer

Hidden layer

Turn $t$

Delexicalised CNN

1\textsuperscript{st} conv.

2\textsuperscript{nd} conv.

3\textsuperscript{rd} conv.

max-pool

avg-pool

Sentence representation
I want Korean food.

Jordan RNN-CNN belief trackers

British French Korean ... Chinese

Output layer

Hidden layer

Turn t

Input layer

Delexicalised CNN
Jordan RNN-CNN belief trackers

British  French  Korean  ...  Chinese

Output layer
Hidden layer
Turn t
Input layer

Delexicalised CNN
Jordan RNN-CNN belief trackers

British  French  Korean ... Chinese

Output layer

Hidden layer

Turn t
Input layer

Delexicalised CNN

Memorise the delex. position
Jordan RNN-CNN belief trackers

British  French Korean ...  Chinese

Output layer

Hidden layer

Turn t
Input layer

Delexicalised CNN

1\textsuperscript{st} conv.
Jordan RNN-CNN belief trackers

British  French  Korean  ...  Chinese

Output layer

Hidden layer

Turn t

Input layer

Pad zeros to have the same length

Delexicalised CNN  1st conv.
Jordan RNN-CNN belief trackers

British  French  Korean  ...  Chinese

Output layer

Hidden layer

Turn t
Input layer

Delexicalised CNN

1$^{\text{st}}$ conv.  2$^{\text{nd}}$ conv.
Jordan RNN-CNN belief trackers

British French Korean ... Chinese

Delexicalised CNN

Output layer

Hidden layer

Turn t

Input layer

I want v.food s.food
Jordan RNN-CNN belief trackers

British  French  Korean ...  Chinese

Output layer

Hidden layer

Turn t

Input layer

Delexicalised CNN  1\textsuperscript{st} conv.  2\textsuperscript{nd} conv.  3\textsuperscript{rd} conv.  max-pool
Jordan RNN-CNN belief trackers

British  French  Korean ... Chinese

Output layer

Hidden layer

Turn t

Input layer

Delexicalised CNN  1\(^{st}\) conv.  2\(^{nd}\) conv.  3\(^{rd}\) conv.  max-pool  avg-pool

sentence representation
Jordan RNN-CNN belief trackers

Delexicalised CNN

1st conv.  2nd conv.  3rd conv.  max-pool  avg-pool

Input layer

Hidden layer

Output layer

Sentence representation

British  French Korean ... Chinese

1.3
Jordan RNN-CNN belief trackers

British  French  Korean ...  Chinese

1.3

Slot-specific delex. ngram feature

Delexicalised CNN  1\textsuperscript{st} conv.  2\textsuperscript{nd} conv.  3\textsuperscript{rd} conv.  max-pool  avg-pool

Input layer

Hidden layer

Output layer

Turn \( t \)

Sentence representation

1.3
Jordan RNN-CNN belief trackers

Delexicalised CNN

1st conv.  2nd conv.  3rd conv.  max-pool  avg-pool

British  French  Korean ...  Chinese

I want v.food s.food

<nil>

Value-specific delex. ngram placeholder

Slot-specific delex. ngram feature

Output layer

Hidden layer

Turn t

Input layer
Jordan RNN-CNN belief trackers

British  French Korean ... Chinese
1.3   2.3

Delexicalised CNN

1st conv. 2nd conv. 3rd conv. max-pool avg-pool
Jordan RNN-CNN belief trackers

Delexicalised CNN

1st conv.
2nd conv.
3rd conv.
max-pool
avg-pool

Output layer
Hidden layer
Turn t
Input layer

British  French  Korean  ...  Chinese
1.3  2.3  9.7

I want v.food s.food

sentence representation
Jordan RNN-CNN belief trackers

British  French  Korean  ...  Chinese
1.3    2.3    9.7

Value-specific delex. ngram feature

Delexicalised CNN  1\textsuperscript{st} conv.  2\textsuperscript{nd} conv.  3\textsuperscript{rd} conv.  max-pool  avg-pool
Jordan RNN-CNN belief trackers

British  French  Korean ...  Chinese
1.3  2.3  9.7  1.2

Delexicalised CNN

1\textsuperscript{st} conv.  2\textsuperscript{nd} conv.  3\textsuperscript{rd} conv.  max-pool  avg-pool

Output layer
Hidden layer
Turn t
Input layer

sentence representation
Jordan RNN-CNN belief trackers

British  French  Korean ...  Chinese
0.01  0.02  0.85  0.01

Input layer

Output layer

Hidden layer

Turn t

Delexicalised CNN

1st conv.  2nd conv.  3rd conv.  max-pool  avg-pool

v.food
s.food

sentence representation
Jordan RNN-CNN belief trackers

British French Korean ... Chinese

0.01 0.02 0.85 0.01

[Henderson et al, 2014]