



UNIVERSITY OF
CAMBRIDGE

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

Google HQ, 23/06/2016
Tsung-Hsien (Shawn) Wen

Dialogue Systems Group

Outline

2

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

Outline

3

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

The two paradigms

4

M: Hello, welcome to the Cambridge dialogue system, What kind of food would you like?
H: Yeah I want to find a restaurant that serves European food.
M: Hotel du Vin and Bistro is a nice place. It serves European food.
H: Uh what is the address and phone number?
M: Hotel du Vin and Bistro is a nice place. Their address is XXX and their phone is YYY.
H: Thank you goodbye.
M: Thank you for using this system.

[Su et al, 2016]

A: How old are you ?
B: I'm 16. Why are you asking ?
A: I thought you were 12.
B: What made you think so ?
A: I don't know what you are talking about.
B: You don't know what you are saying.

[Li et al, 2016]

The two paradigms

5

- ⊙ Task-Oriented Dialogue Systems
 - ⊙ Goal-oriented
 - ⊙ Require precise understanding, hard to collect data.
 - ⊙ Modular, highly handcrafted, restricted ability, but meaningful/useful systems.
- ⊙ Chat-based Conversational Agents
 - ⊙ Chit-chat (non-goal).
 - ⊙ Vast amount of data (but probably not helpful).
 - ⊙ End-to-end, highly data-driven, but meaningless/inappropriate responses, unreliable systems.
- ⊙ Can we train a useful (complete tasks) dialogue system directly from data?
- ⊙ How can we collect the data to train this model?

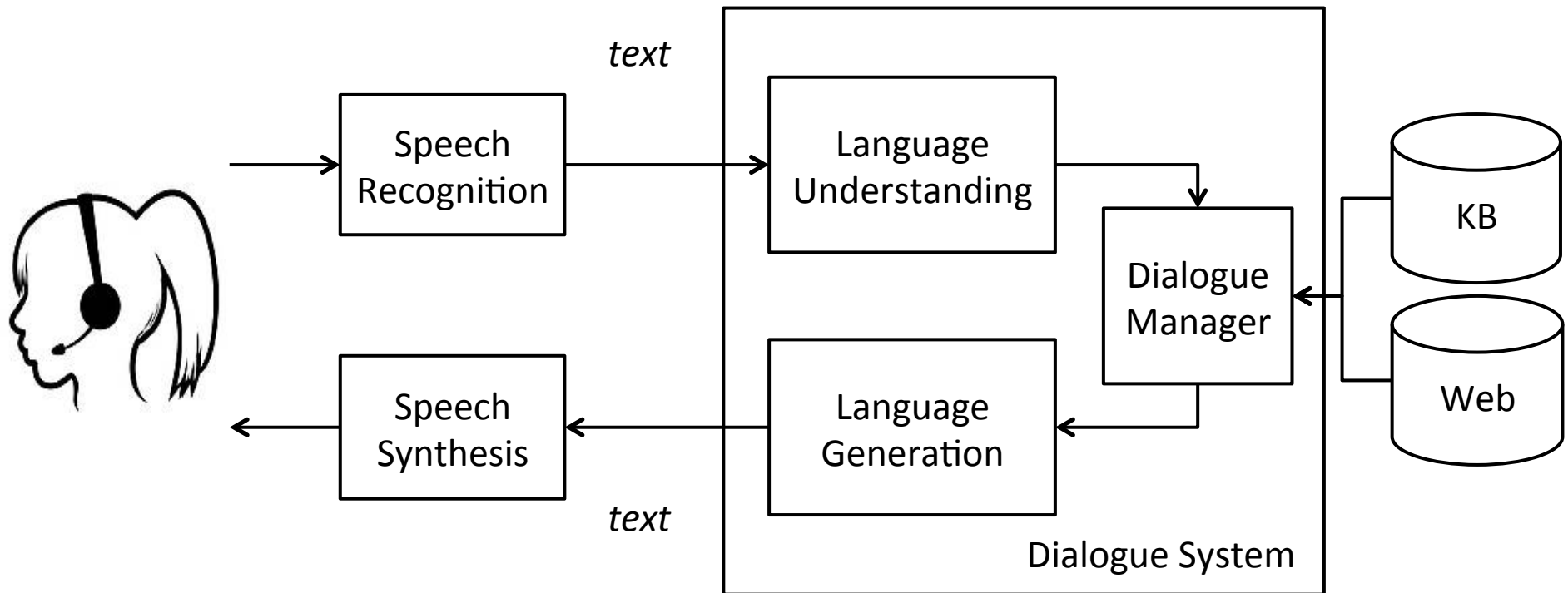
Outline

6

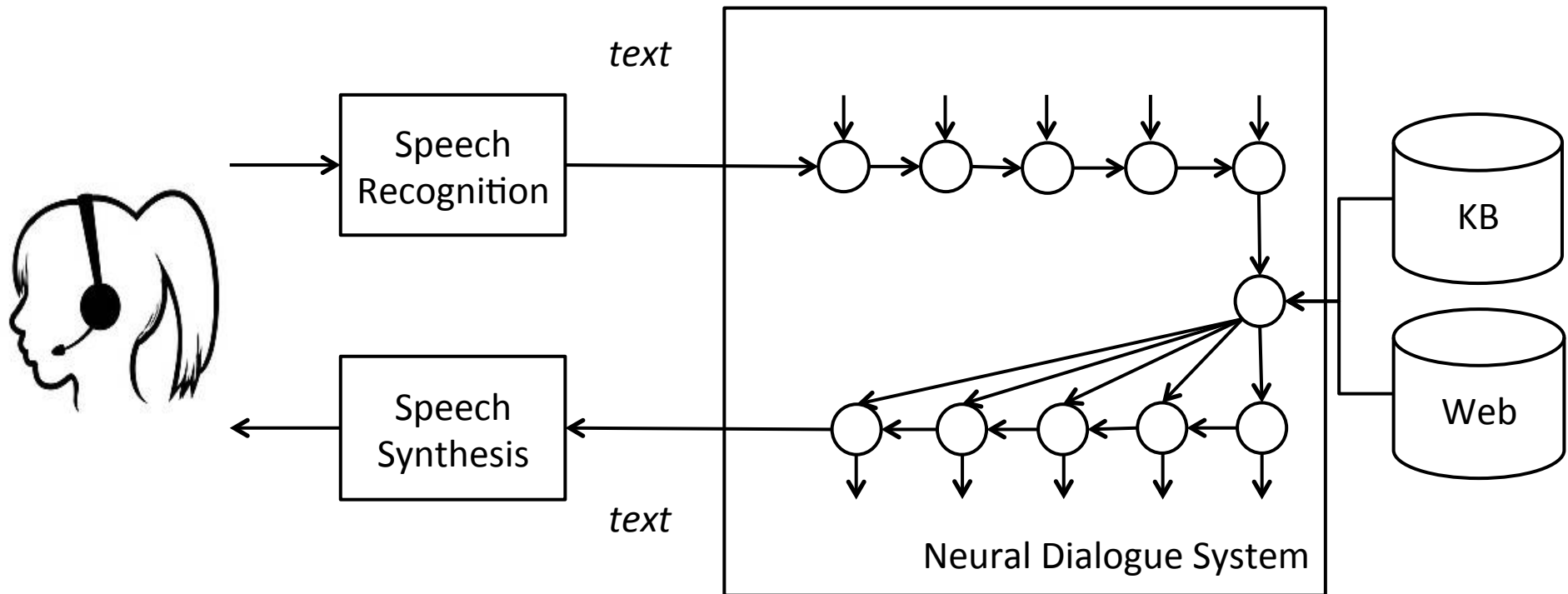
- ⊙ Intro
- ⊙ Neural Dialogue System
- ⊙ Wizard-of-Oz Data Collection
- ⊙ Experiments
- ⊙ Conclusion & Discussion

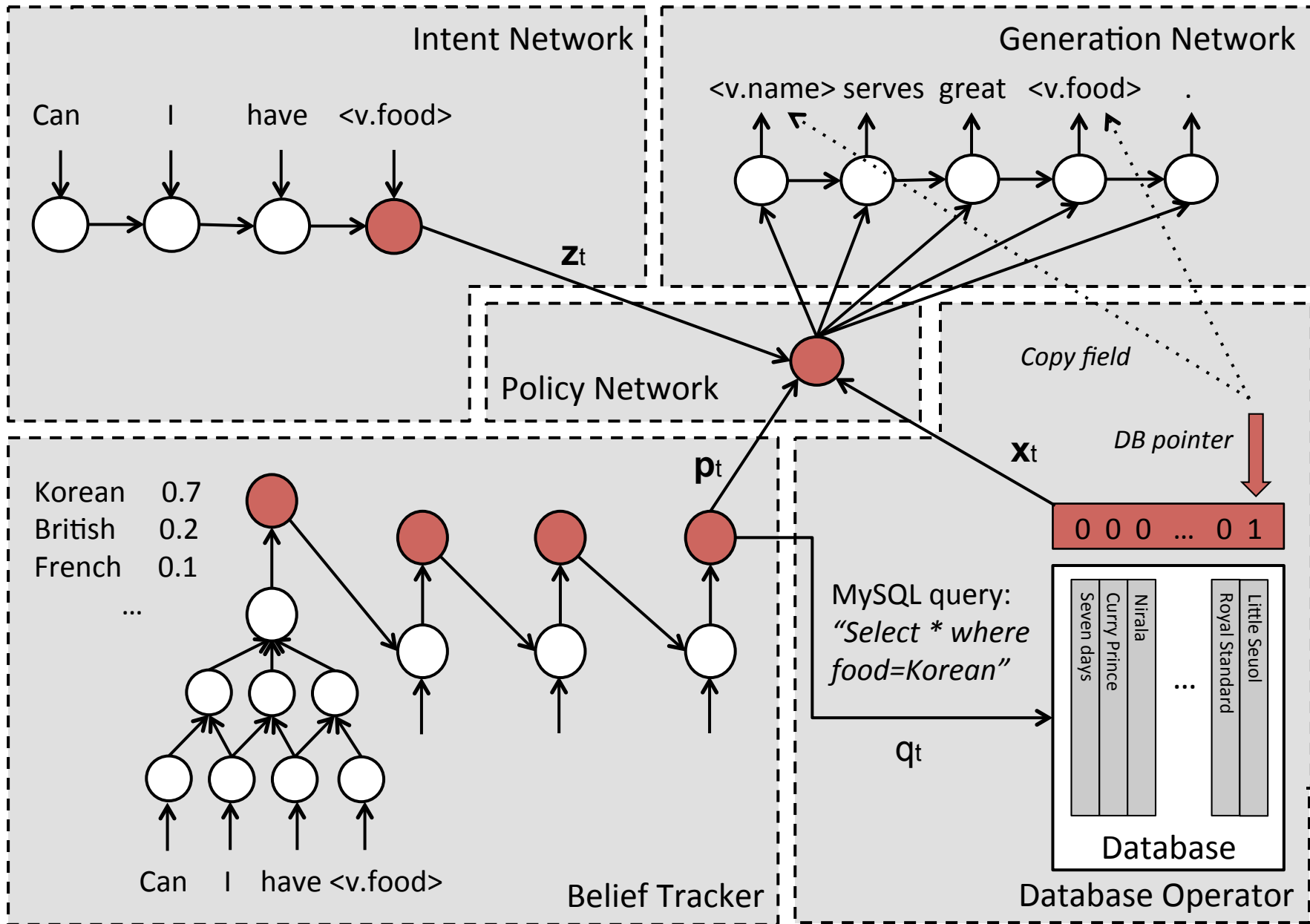
Traditional Dialogue Systems

7



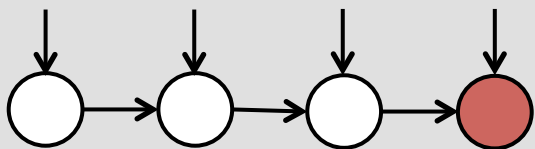
Neural Dialogue Systems

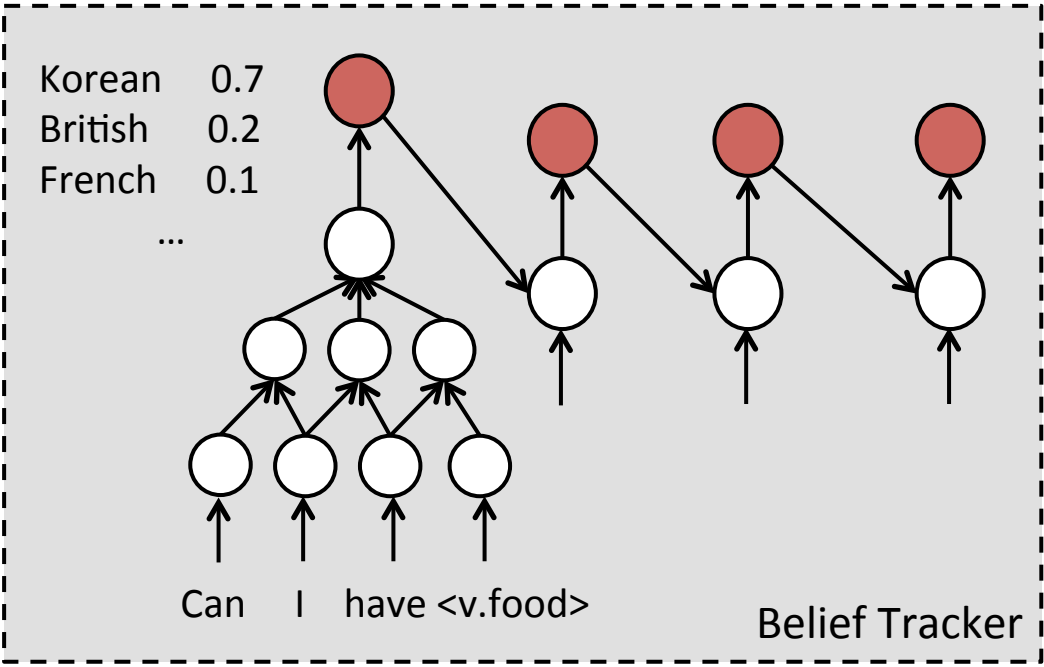
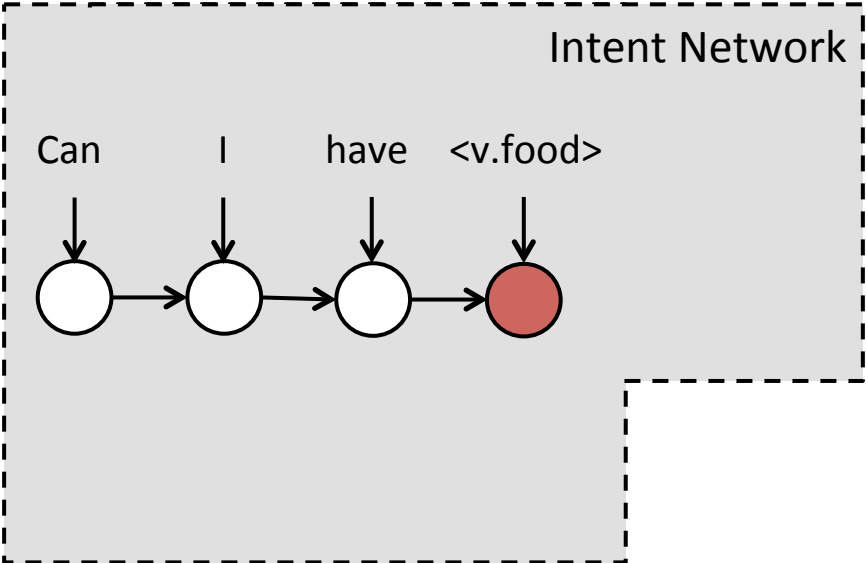




Intent Network

Can I have <v.food>





Jordan RNN-CNN belief trackers

12

British French Korean ... Chinese

.03 .03 .857 ... 102

[Henderson et al, 2014]

Output layer

Hidden layer

Turn t
Input layer

Value-specific
delex. ngram
placeholder

Slot
delex

Value-specific
delex. ngram
feature

Pad zeros to
have the
same length

Memorise
the delex.
position

sentence
representation

→ Korean
→ food

Delexicalised CNN

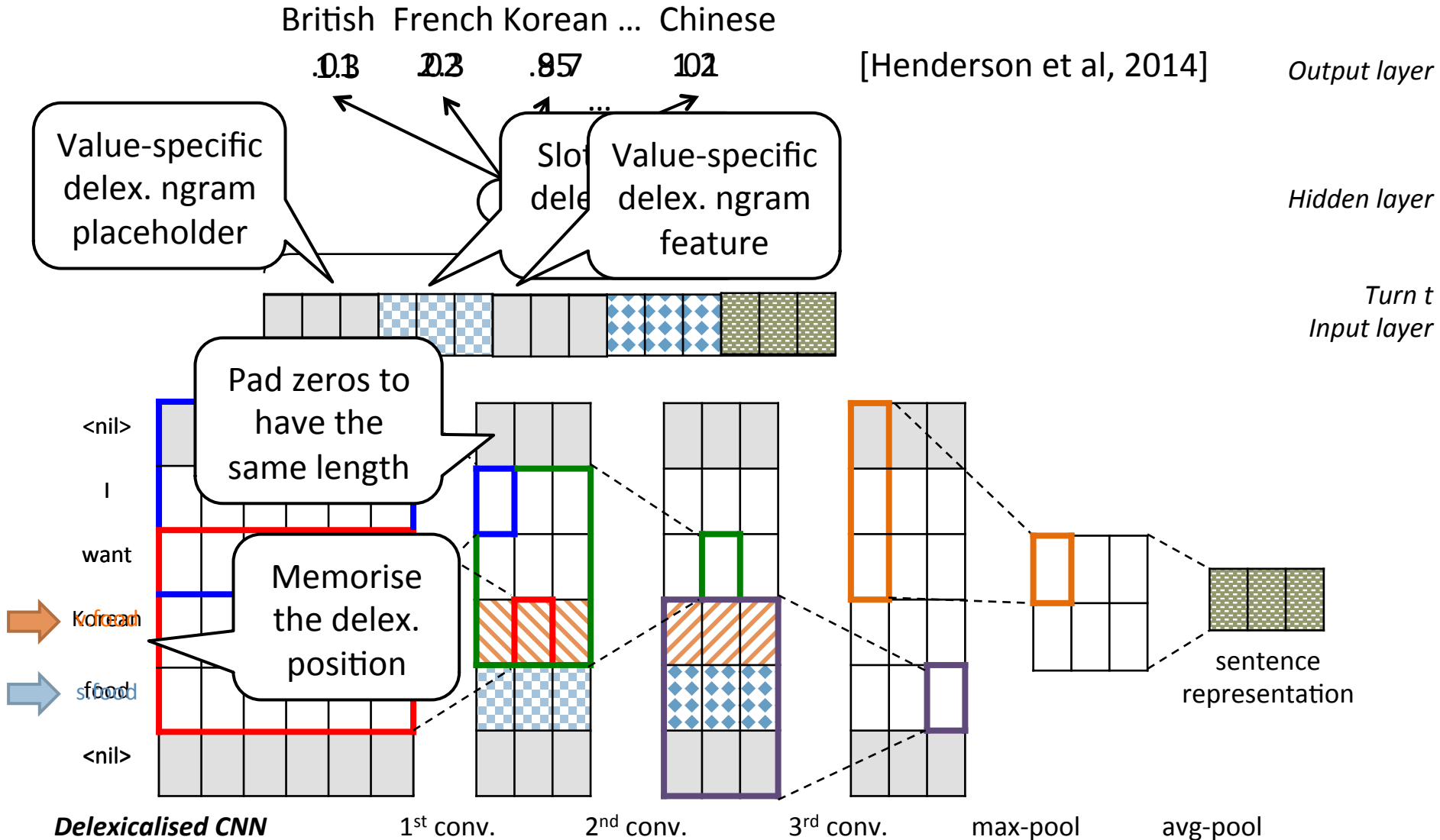
1st conv.

2nd conv.

3rd conv.

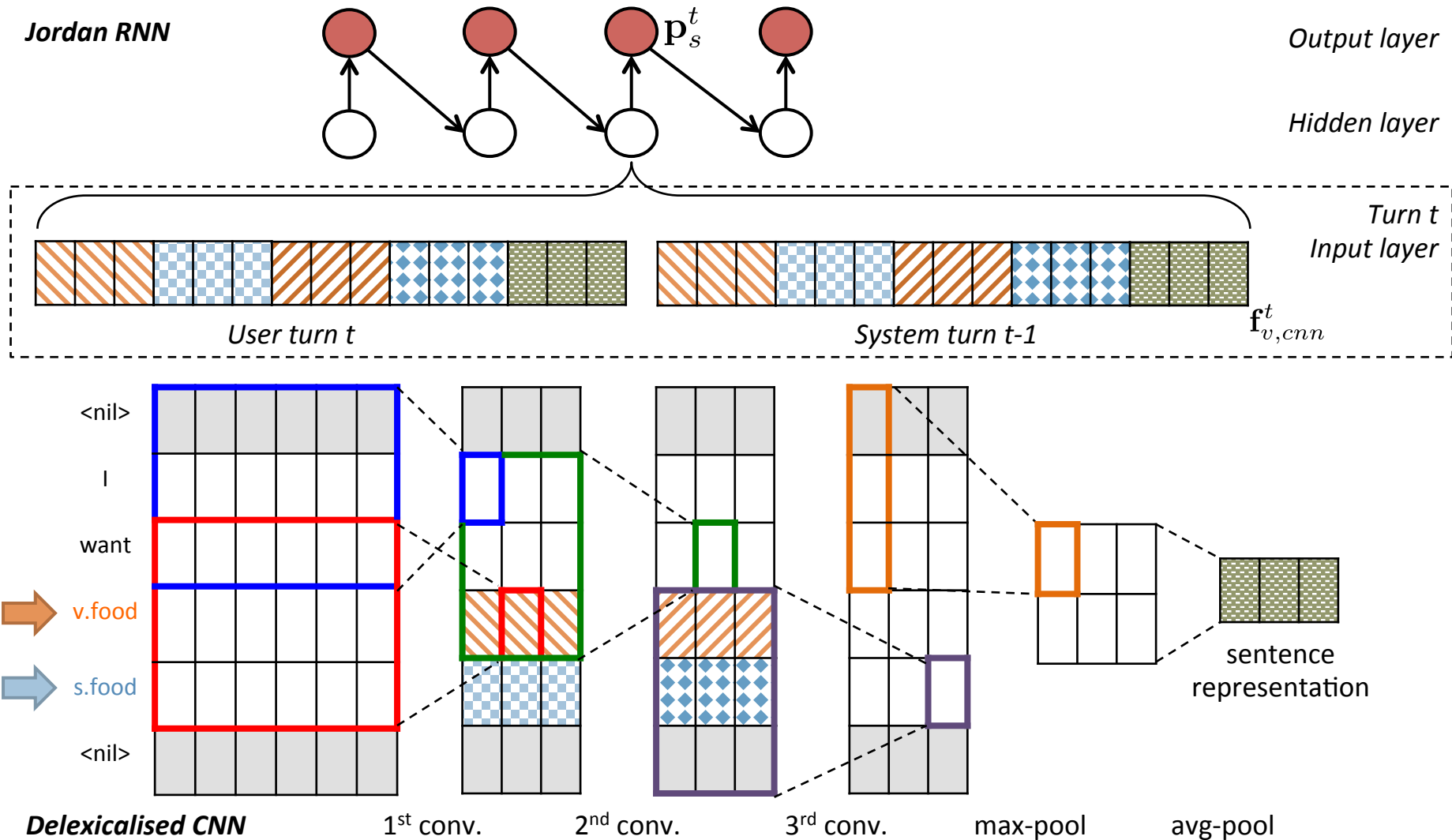
max-pool

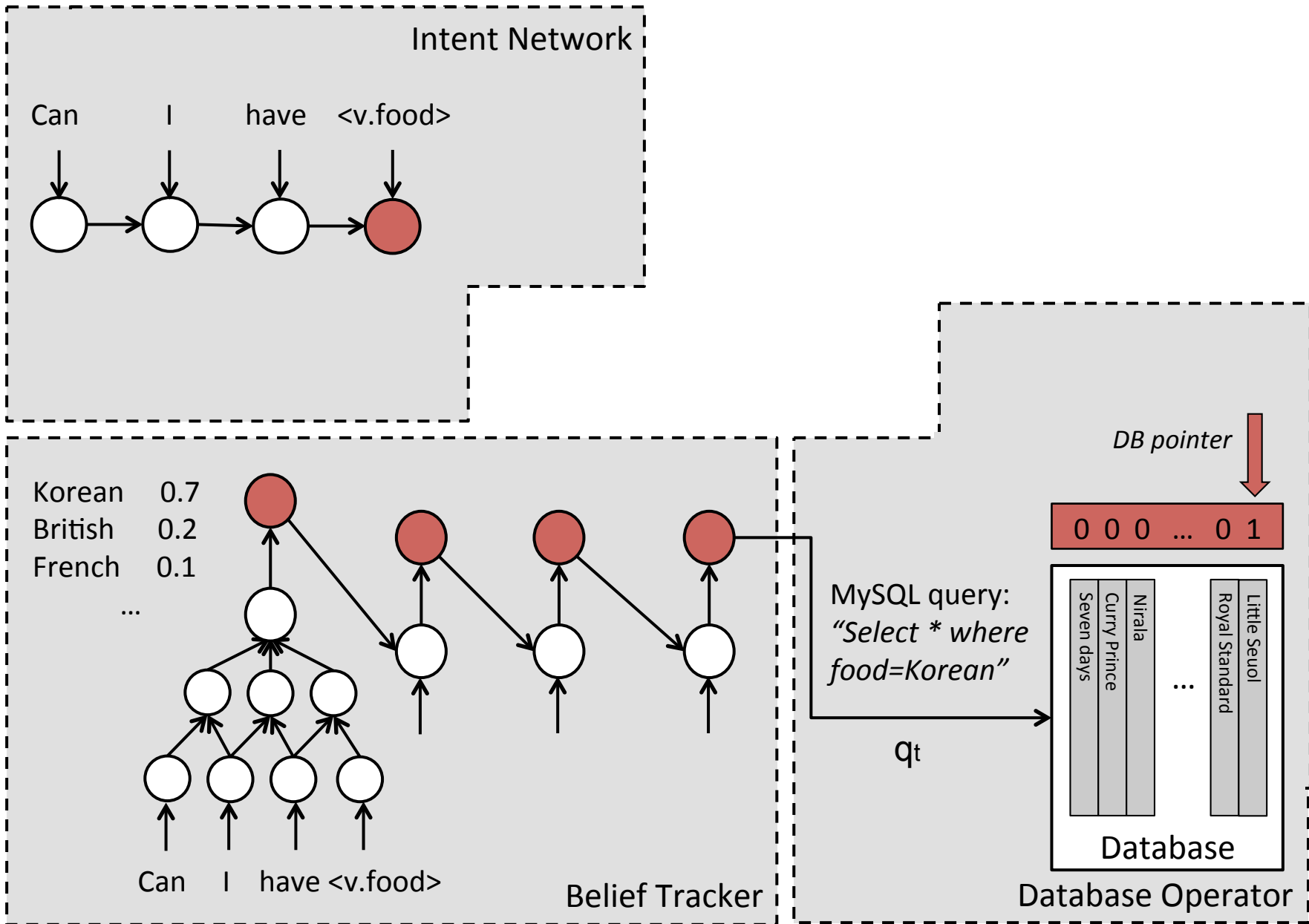
avg-pool



Jordan RNN-CNN belief trackers

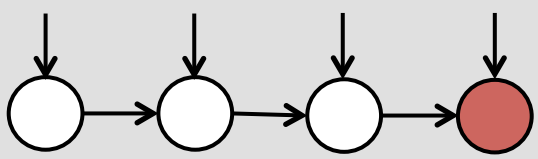
13



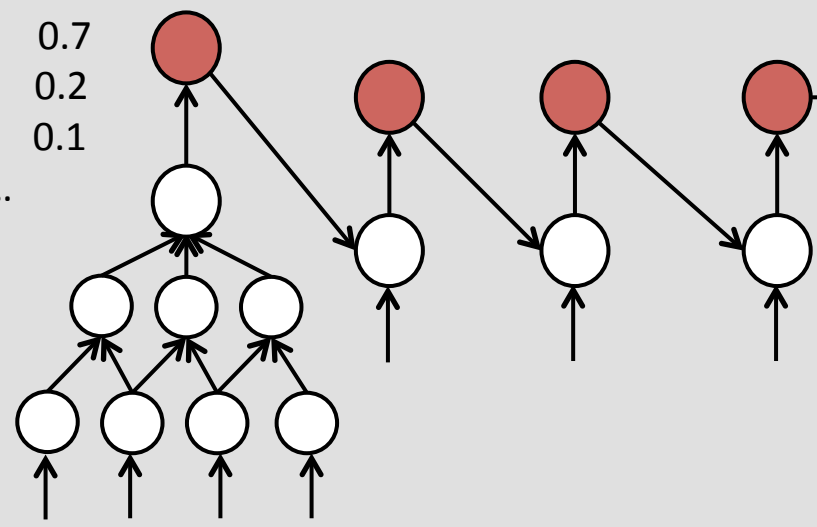


Intent Network

Can I have <v.food>



Korean 0.7
 British 0.2
 French 0.1
 ...



Can I have <v.food>

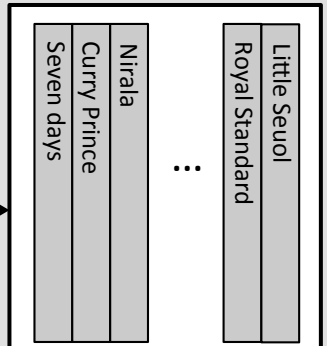
Belief Tracker

DB pointer

0 0 0 ... 0 1

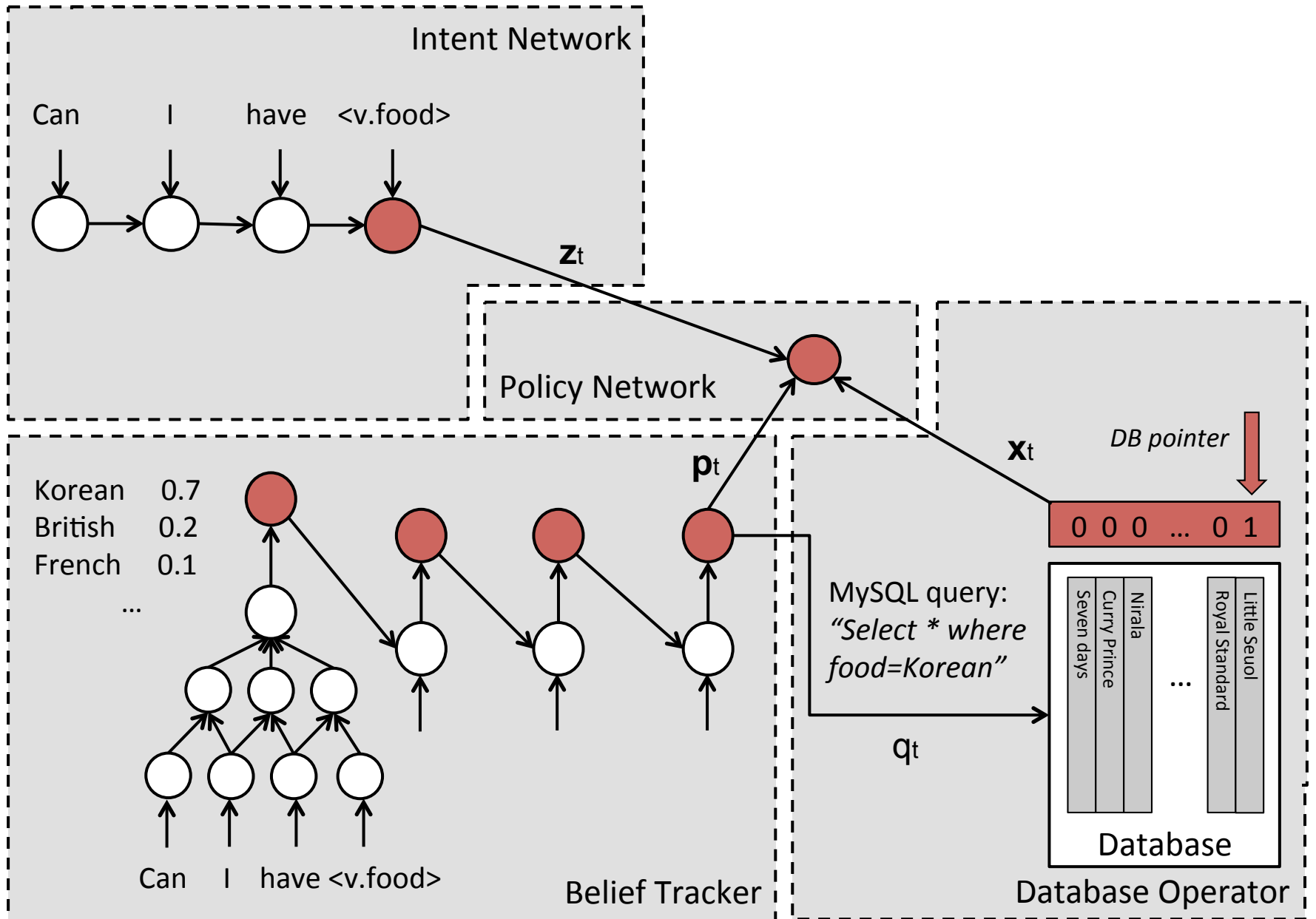
MySQL query: "Select * where food=Korean"

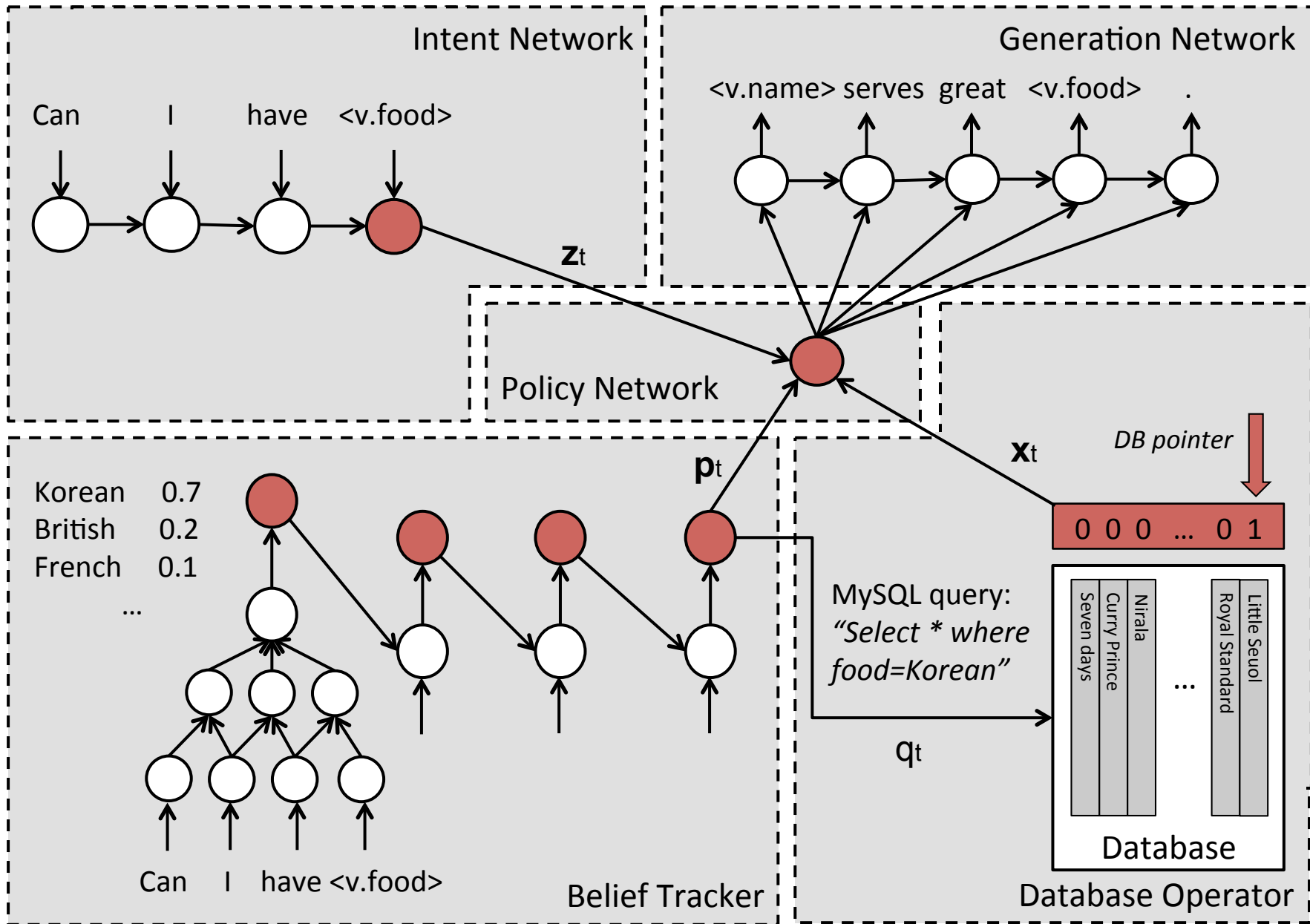
qt

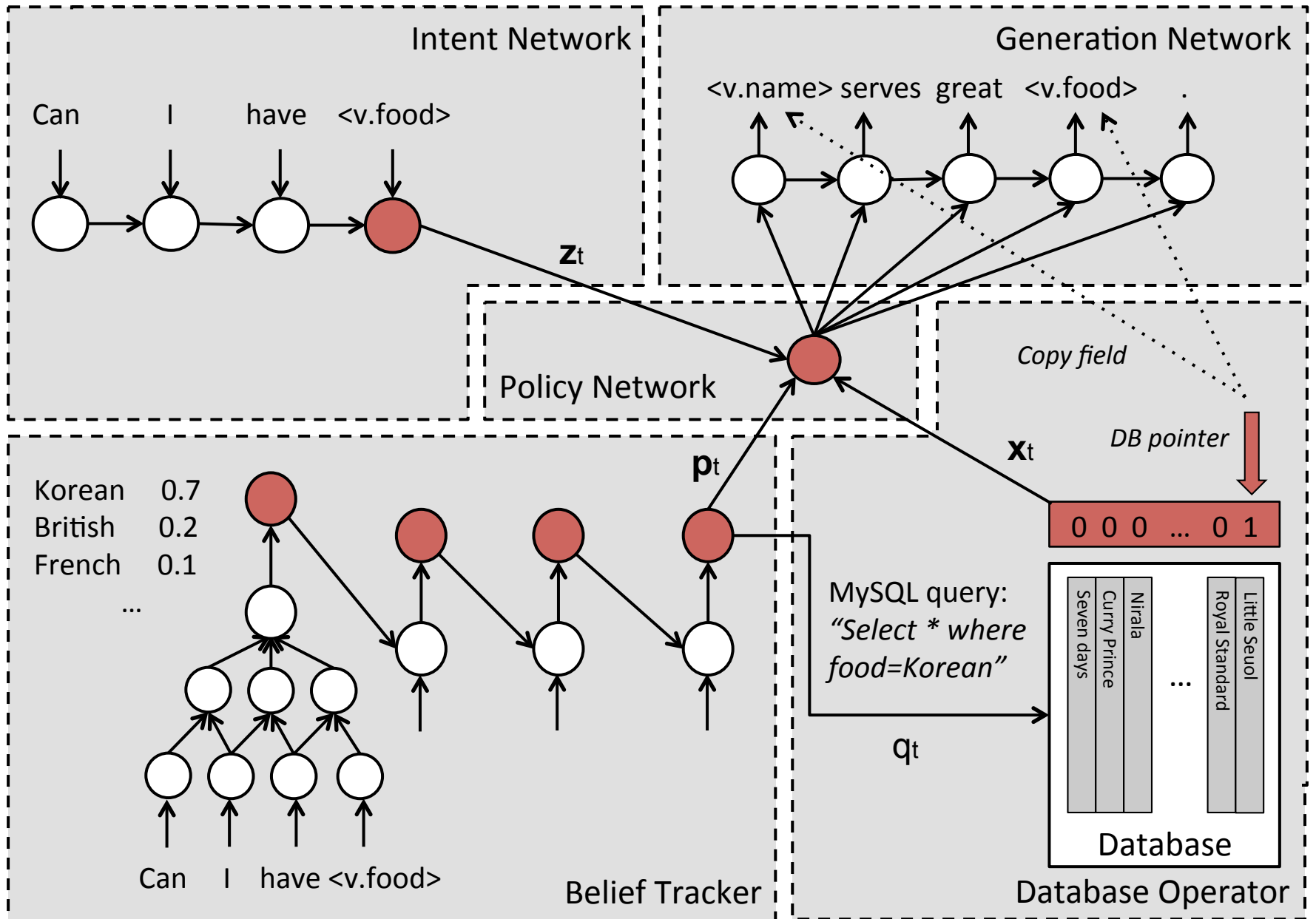


Database

Database Operator







Outline

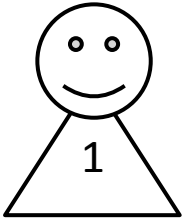
18

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

Wizard of Oz Data Collection

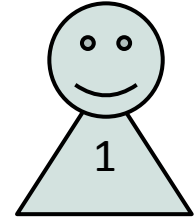
19

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?

Food	Korean
Price	Cheap
Area	N/A

Search Table

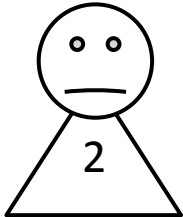
Little Seoul	...
Best Korea	...

...

Wizard of Oz Data Collection

20

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

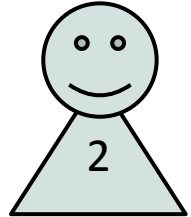


Hi, I want a cheap Korean restaurant.

What area are you looking for ?

Somewhere in the north.

Little Seoul is nice one in the north.



What user wants?

Food	Korean
Price	Cheap
Area	North

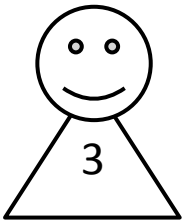
Search Table

Little Seoul	...
--------------	-----

Wizard of Oz Data Collection

21

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



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

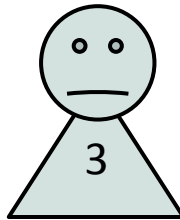
*What **area** are you looking for ?*

*Somewhere in the **north**.*

Little Seoul is nice one in the north.

*What is the **phone number**?*

*Its phone number is **01223456789**.*



What user wants?

Food	Korean
Price	Cheap
Area	North

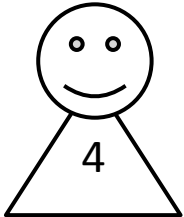
Search Table

Little Seoul	...
--------------	-----

Wizard of Oz Data Collection

22

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**.*

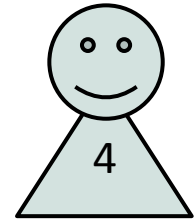
Little Seoul is nice one in the north.

*What is the **phone number**?*

*Its phone number is **01223456789**.*

Thank you very much, good bye.

Thank you for using the system.



What user wants?

Food	Korean
Price	Cheap
Area	North

Search Table

Little Seoul	...
--------------	-----

Wizard of Oz Data Collection

23

*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.

What user wants?

Food	Korean
Price	Cheap
Area	North

Wizard of Oz Data Collection

24

- ⦿ 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.
- ⦿ Example user page

Task 02004: You are looking for and it should serve **gastropub food**. You don't care about the **price range**. You want to know the **address**.

Info Desk : Hello , welcome to the Cambridge restaurant system . You can ask for restaurants by area, price range or food type . How may I help you ?

Customer : i want a gastropub food

Info Desk : There are 4 restaurants serving gastropub food, what price range do you want ?

Next turn!

Customer : (Your response)

i don't care about the price range, just give me the address please.

Submit the HIT

Wizard of Oz Data Collection

25

Example wizard page

Info Desk : Hello , welcome to the Cambridge restaurant system . You can ask for restaurants by area, price range or food type . How may I help you ?
Customer : i want a gastropub food
Info Desk : There are 4 restaurants serving gastropub food, what price range do you want ?
Customer : i dont care

Next turn!

Please **modify** the following answers based on the latest customer response:

- What does user want?

What is the **food type** the user wants?

What is the **area** the user wants?

What is the **price range** the user wants?

- What does user ask?

Is the user asking for **food type** of an offered venue?

Is the user asking for **price range** of an offered venue?

Is the user asking for **area** of an offered venue?

Is the user asking for **postcode** of an offered venue?

Is the user asking for **phone number** of an offered venue?

Is the user asking for **address** of an offered venue?

Is the user mentioning any **restaurant names**?

Info Desk : (Your response)

I would recommend backstreet bistro, a great gastropub restaurant in the centre. do you want their phone number ? end-of-dialogue?

Name	Food	Area	Price Range	Phone	Address	Postcode
backstreet bistro	gastropub	centre	expensive	01223 306306	2 Sturton Street City Centre	C.B 1, 2 Q.A
royal standard	gastropub	east	expensive	01223 247877	290 Mill Road City Centre	C.B 1, 3 N.L
the cow pizza kitchen and bar	gastropub	centre	moderate	01223 308871	Corn Exchange Street	C.B 2, 3 Q.F
the slug and lettuce	gastropub	centre	expensive	--	34 - 35 Green Street	C.B 2, 3 J.U
nil	gastropub	nil	nil	nil	nil	nil

Showing 1 to 4 of 4 entries (filtered from 110 total entries)

Previous Next

Data Statistics

26

- ⊙ 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

27

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

Experiments

28

- ⊙ 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

Response Generation Task

29

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

Human evaluation

30

Quality assessment

Metric	N2N
Success	98%
Comprehension	4.11
Naturalness	4.05
# of dialogues:	245

System Comparison

Metric	N2N	Modular	Tie
Subj. Success	96.95%	95.12%	-
Avg. # of Turn	3.95	4.54	-
Comparisons(%)			
Naturalness	46.95*	25.61	27.44
Comprehension	45.12*	21.95	32.93
Preference	50.00*	24.39	25.61
Performance	43.90*	25.61	30.49

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

Example dialogues

31

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 .

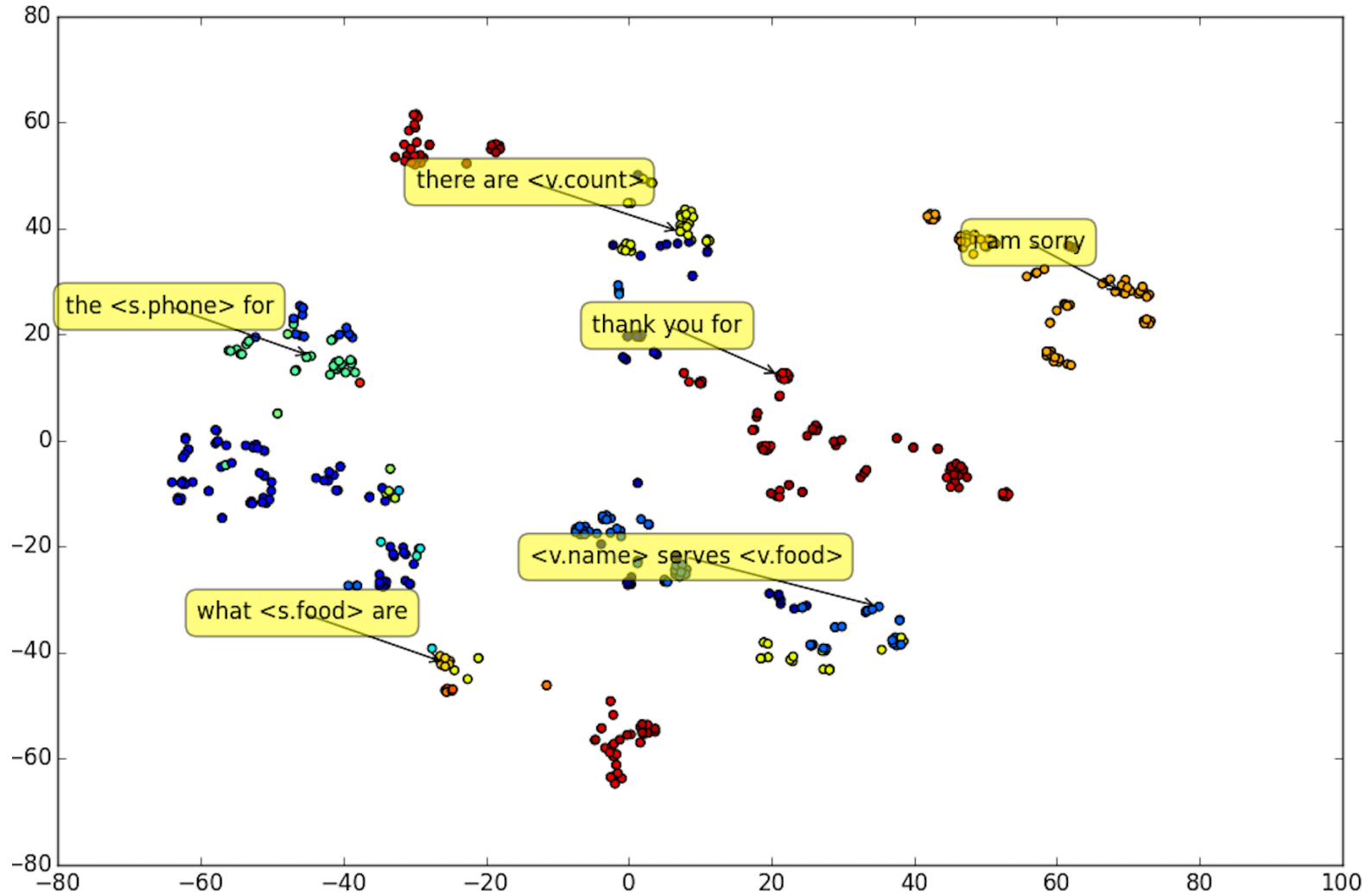
Example dialogues

32

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 .

Visualising action embedding

33



Outline

34

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

Conclusion & Discussion

35

- ⊙ 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.
- ⊙ Explicit language grounding is crucial, but what is the best way to represent semantics?

The paper

36

- ⊙ Tsung-Hsien Wen, David Vandyke, Nikola Mrksic, Milica Gasic, Lina M.R. Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young. A Network-based End-to-End Trainable Task-oriented Dialogue System. *arXiv preprint: 1604.04562* 2016.

References

- ⊙ On-line Active Reward Learning for Policy Optimisation in Spoken Dialogue Systems, P-H. Su, M. Gasic, N. Mrksic, L. Rojas-Barahona, S. Ultes, D. Vandyke, T-H. Wen, and S. Young. ACL 2016.
- ⊙ Word-Based Dialog State Tracking with Recurrent Neural Networks, M. Henderson, B. Thomson and S. Young. SigDial 2014.
- ⊙ Deep Reinforcement Learning for Dialogue Generation, J. Li, W. Monroe, A. Ritter, D. Jurafsky. arXiv preprint 1606.01541 2016.



UNIVERSITY OF
CAMBRIDGE

Thank you! Questions?

*Tsung-Hsien Wen is supported by a studentship funded by Toshiba
Research Europe Ltd, Cambridge Research Laboratory*

Dialogue Systems Group