Dialogue State Tracking
as Question Answering

Hung-yi Lee 李宏毅

Dialogue State Tracking (DST)

- It is the core of task-oriented dialogue.

https://dstc9.dstc.community/

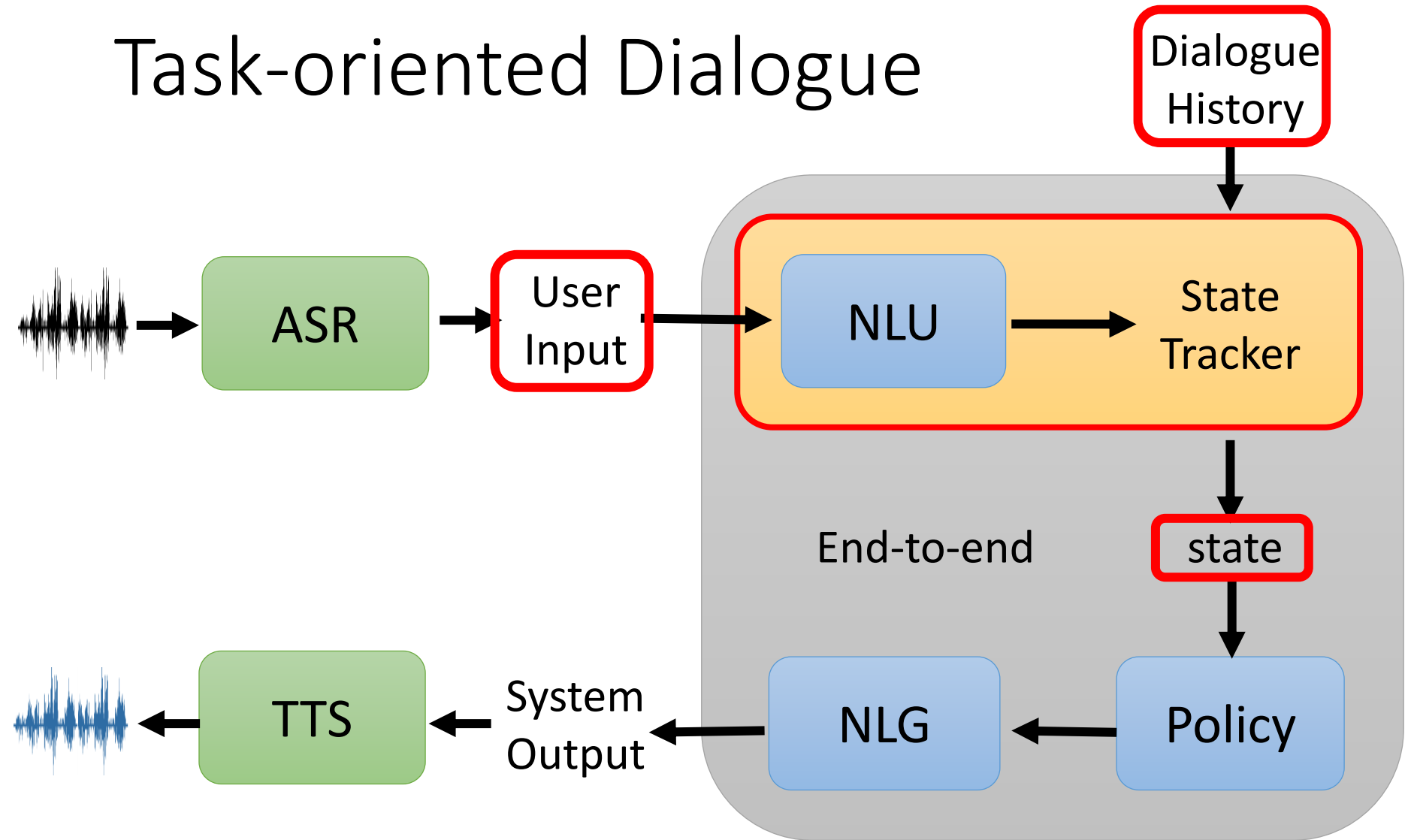It was called *Dialog State Tracking Challenge* (DSTC) before 2017.

- State-of-the-art approaches consider **DST** as **QA**.

Task-oriented Dialogue

ASR → User Input → NLU → State Tracker → state → Policy → NLG → System Output → TTS

Dialogue History

End-to-end

Input: several sequences, Output: a sequence

Dialogue State Tracking

User: I will be traveling to Cambridge and cannot wait to try the restaurants, but could you help me with finding a hotel?

Agent: Would you prefer a guesthouse? the Acorn guest house is in the north and moderately priced. 4 starts with free internet and parking.

User: I would prefer a **hotel type hotel**, it just needs to be in the **moderate price range** and does not need to have free parking, but **needs wifi**.

Agent: Sure! I found **Ashley hotel**. It is located at 74 Chesterton road. Would that work for you?

User: **Yes**, that sounds like a good place. I appreciate the information.

Agent: You are welcome. Would you like a room, or perhaps I could help you with something else?

User: Yes, can you tell me

Agent: It has 2 stars.

User: Thanks, I also need

Agent: How about the *Gal*

User: **Sounds good**. I woul

Agent: Okay I was able to

User: Could you also book

Agent: Where will you be

User: **From the hotel**. I wo

Agent: I was able to book

else today?

User: No, that will be all. T

Source: https://whatscap.com/share/49576

Dialogue State Tracking

User: I will be traveling to Cambridge and cannot wait to try the restaurants, but could you help me with finding a hotel?

Agent: Would you prefer a guesthouse? the Acorn guest house is in the north and moderately priced. 4 starts with free internet and parking.

User: I would prefer a **hotel type hotel**, it just needs to be in the **moderate price range** and does not need to have free parking, but **needs wifi**.

Agent: Sure! I found **Ashley hotel**. It is located at 74 Chesterton road. Would that work for you?

User: **Yes**, that sounds like a good place. I appreciate the information.

Agent: You are welcome. Would you like a room, or perhaps I could help you with something else?

User: Yes, can you tell me the star rating of that hotel?

Agent: It has 2 stars.

User: Thanks, I also need a restaurant in the **centre** that serves **European food**. It needs to be in the _same price range as hotel_.

Agent: How about the **Galleria**?

User: **Sounds good**. I would like to make a reservation for **Friday at 17:15 for 4 people**.

Agent: Okay I was able to make you a reservation and your reference number is DWMSURSO.

User: Could you also book me a taxi that _arrives at the restaurant by the time of my reservation_?

Agent: Where will you be departing from?

User: **From the hotel**. I would like to get [...] something changes.

Agent: I was able to book that taxi for yo[...] will be a yellow Honda. Can I assist you with anything else today?

User: No, that will be all. Thank you, goo[...]

Dialogue State Tracking

hotel type: hotel
hotel price range: moderate
hotel Internet: yes
hotel name: Ashley hotel

restaurant area: centre
restaurant food: European
restaurant price range: moderate
restaurant name: Galleria
restaurant book day: Friday
restaurant book time: 17:15
restaurant book people: 4

taxi departure: Ashley hotel
taxi destination: Galleria
taxi arrive by: 17:15

State

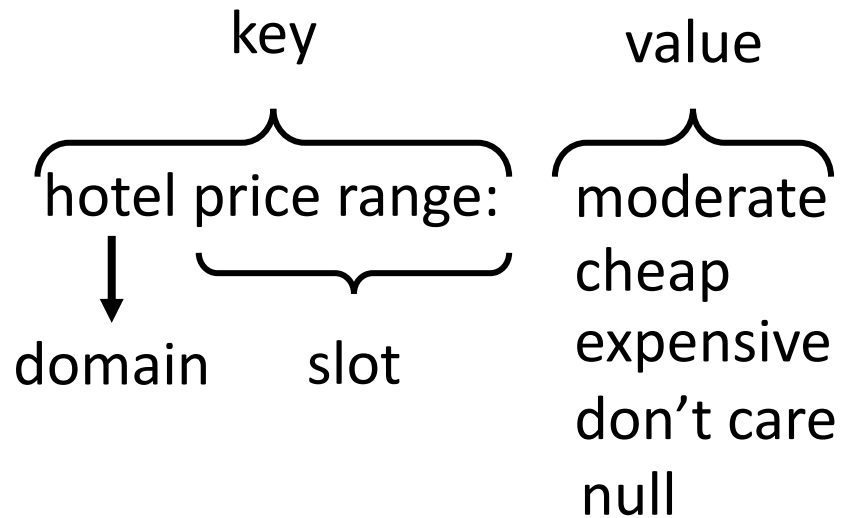Summarization of what has happened in the dialogue (for policy module to make decision)

Dialogue State Tracking (DST)

hotel type: hotel
hotel price range: moderate
hotel Internet: yes
hotel name: Ashley hotel

restaurant area: centre
restaurant food: European
restaurant price range: moderate
restaurant name: Galleria
restaurant book day: Friday
restaurant book time: 17:15
restaurant book people: 4

taxi departure: Ashley hotel
taxi destination: Galleria
taxi arrive by: 17:15

Dialogue State

A set of (key, value)

key   value

hotel price range: moderate
                        cheap
↓   expensive
domain   slot   don't care
                        null
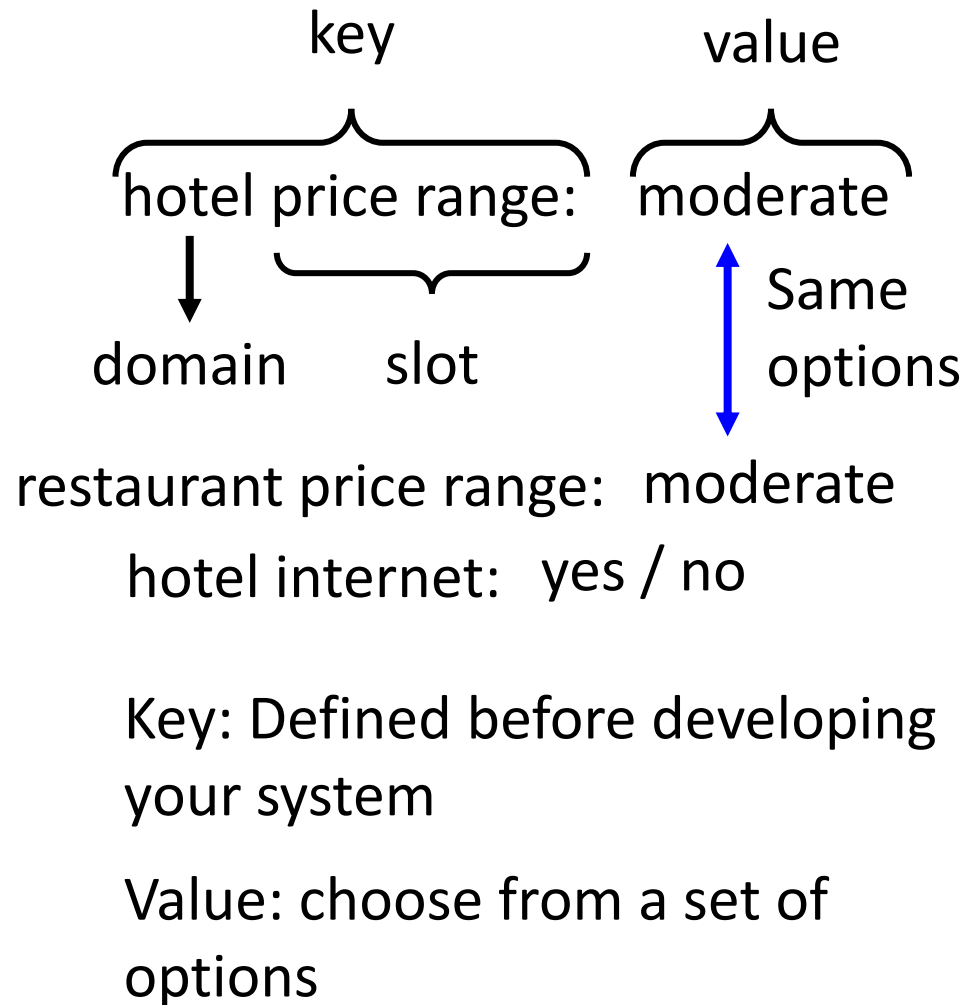
Key: Defined before developing your system

Value: choose from a set of predefined options
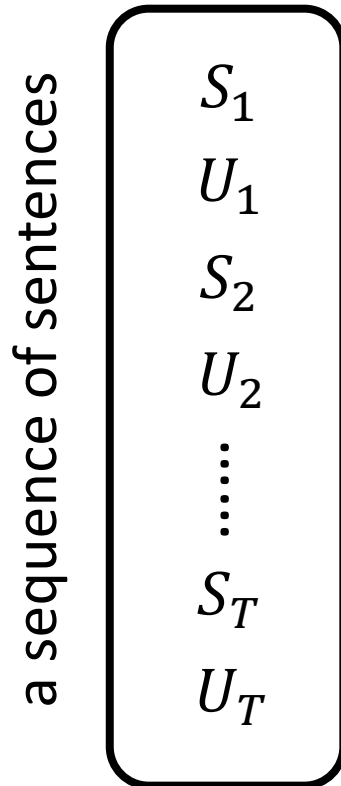
Dialogue State Tracking (DST)

hotel type: hotel
hotel price range: moderate
hotel Internet: yes
hotel name: Ashley hotel

restaurant area: centre
restaurant food: European
restaurant price range: moderate
restaurant name: Galleria
restaurant book day: Friday
restaurant book time: 17:15
restaurant book people: 4

taxi departure: Ashley hotel
taxi destination: Galleria
taxi arrive by: 17:15

Dialogue State

A set of (key, value)

key   value

hotel price range: moderate

domain   slot   Same options

restaurant price range: moderate

hotel internet: yes / no

Key: Defined before developing your system

Value: choose from a set of options

Dialogue State Tracking (DST)

a sequence of sentences

$$S_1$$
$$U_1$$
$$S_2$$
$$U_2$$
$$\vdots$$
$$S_T$$
$$U_T$$

Key k include domain and slot

$$k_1: \quad \{v_1^1, \boxed{v_2^1}, \dots, v_{n_1}^1\}$$

$$k_2: \quad \{\boxed{v_1^2}, v_2^2, \dots, v_{n_2}^2\}$$

$$k_3: \quad \{v_1^3, v_2^3, \dots, \boxed{v_{n_3}^3}\}$$

Select the correct value for each key

What the user and agent have said

S: agent, U: user

MultiWOZ 2.0 [Budzianowski, et al., EMNLP'18]

- Multi-domain   There are 7 domains actually .

	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

What is WOZ?

DIGITALLY RESTORED · REMASTERED IN DIGITAL STEREO SOUND
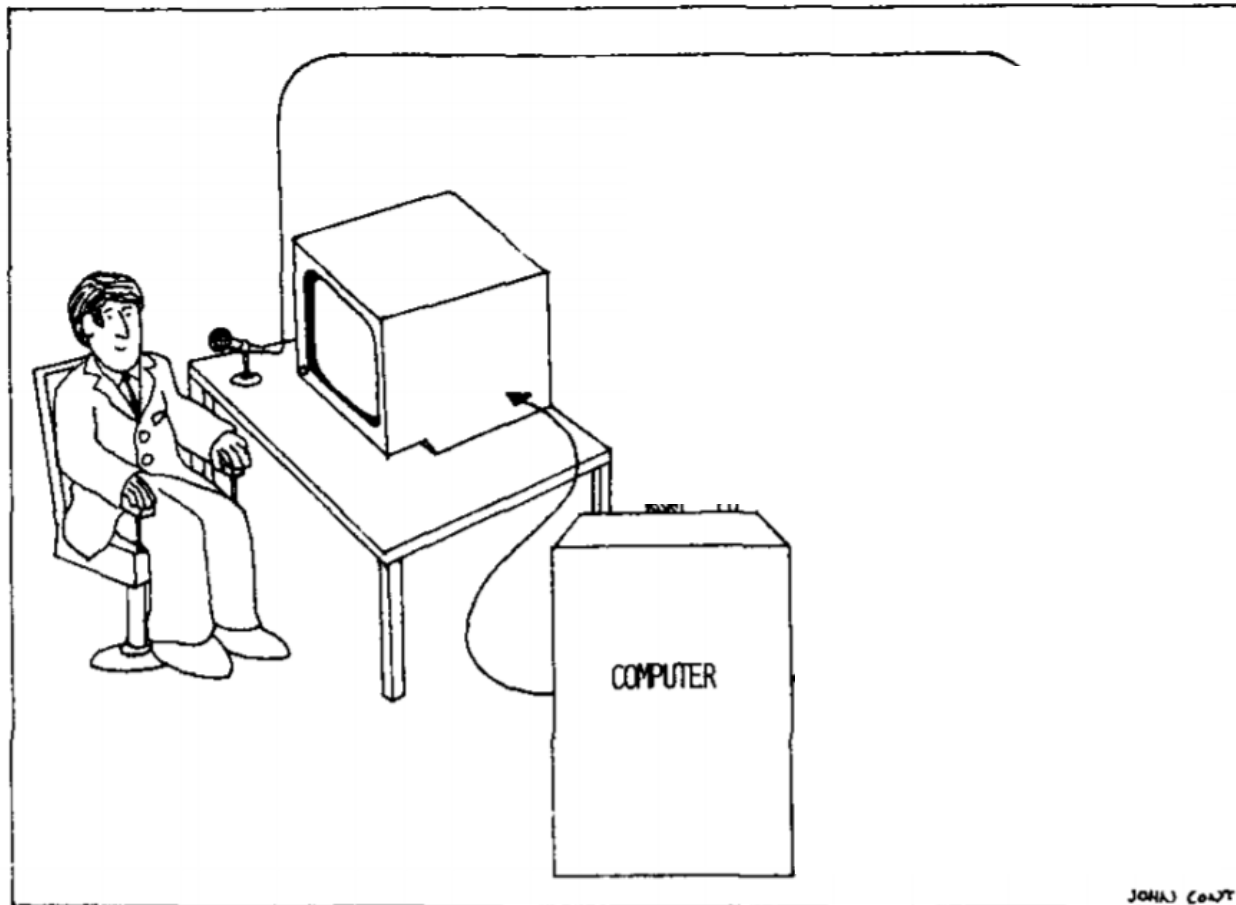SPECIAL EDITION
THE WIZARD OZ

MakeAGIF.com

What is WOZ?

- One way of system test / data collection

Artificial AI (AAI)
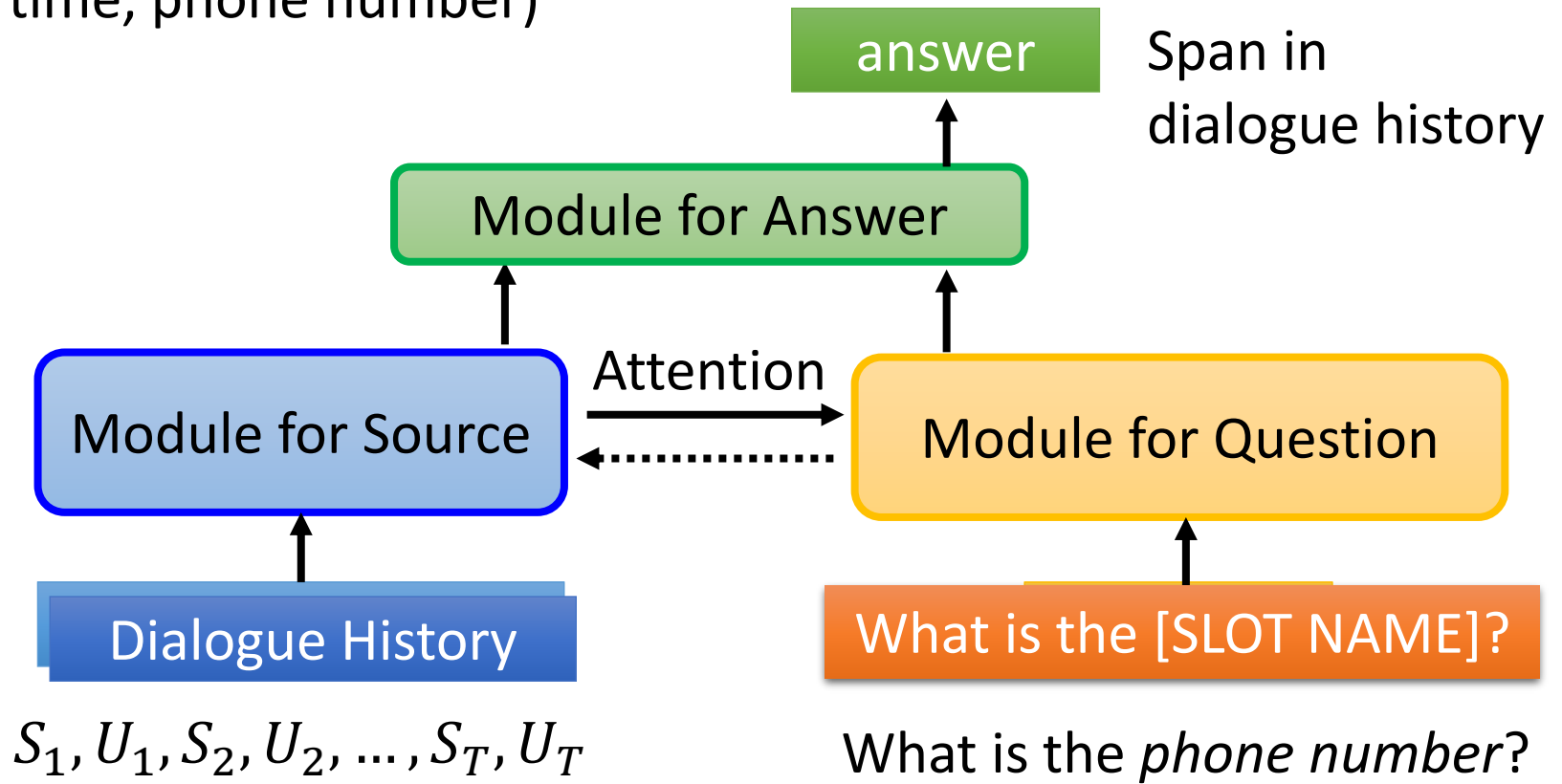
Schema-Guided Dialogue (SGD)

[Rastogi, et al., AAAI'20]

Metric ↓ Dataset →	DSTC2	WOZ2.0	FRAMES	M2M	MultiWOZ	SGD
No. of domains	1	1	3	2	7	**16**
No. of dialogues	1,612	600	1,369	1,500	8,438	**16,142**
Total no. of turns	23,354	4,472	19,986	14,796	113,556	**329,964**
Avg. turns per dialogue	14.49	7.45	14.60	9.86	13.46	**20.44**
Avg. tokens per turn	8.54	11.24	12.60	8.24	**13.13**	9.75
Total unique tokens	986	2,142	12,043	1,008	23,689	**30,352**
No. of slots	8	4	61	13	24	**214**
No. of slot values	212	99	3,871	138	4,510	**14,139**

REQUEST(location)
REQUEST(cuisine)

INFORM(location="los angeles")
INFORM(date="2019-03-02")
INFORM(cuisine="greek")

OFFER(restaurant="Opa!")

→

Which city are you in?
What kind of food?

I want to eat in LA.
I want to go there tomorrow.
I want to eat greek food.

Opa is a nice restaurant.

→

Where are you located and
what is your preferred food?

I am looking for a greek
restaurant in LA for tomorrow.

How does Opa sound?

simulator   template   human

CrossWOZ [Zhu, et al., TACL'20]

Type	Single-domain goal					Multi-domain goal		
Dataset	DSTC2	WOZ 2.0	Frames	KVRET	M2M	MultiWOZ	Schema	**CrossWOZ**
Language	EN	EN	EN	EN	EN	EN	EN	**CN**
Speakers	H2M	H2H	H2H	H2H	M2M	H2H	M2M	**H2H**
# Domains	1	1	1	3	2	7	16	**5**
# Dialogues	1,612	600	1,369	2,425	1,500	8,438	16,142	**5,012**
# Turns	23,354	4,472	19,986	12,732	14,796	115,424	329,964	**84,692**
Avg. domains	1	1	1	1	1	1.80	1.84	**3.24**
Avg. turns	14.5	7.5	14.6	5.3	9.9	13.7	20.4	**16.9**
# Slots								**72**
# Values								**7,871**

CrossWOZ

usr: 你好，可以帮我推荐一个评分是4.5分以上的景点吗？
Hello, could you recommend an attraction with a rating of 4.5 or higher?
sys: 天安门城楼，簋街小吃和北京欢乐谷都是很不错的地方呢。
Tiananmen, Gui Street, and Beijing Happy Valley are very nice places.
usr: 我喜欢北京欢乐谷，你知道这个景点周边的酒店都是什么吗？
I like Beijing Happy Valley. What hotels are around this attraction?
sys: 那可多了，有A酒店, B酒店, C酒店。
There are many, such as hotel A, hotel B, and hotel C.
usr: 太好了，我正打算在**景点**附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？
Great! I am planning to find a hotel to stay **near the attraction**. Which one has a rating of 4 or higher and offers wake-up call service?

Challenge - Infinite values

- Some slots have almost infinite possible values (e.g. leave time, phone number)

Span in dialogue history

answer

Module for Answer

Module for Source

Attention

Module for Question

Dialogue History

$S_1, U_1, S_2, U_2, ..., S_T, U_T$

What is the [SLOT NAME]?

What is the *phone number*?

Challenge – New Service

Having training data   New service

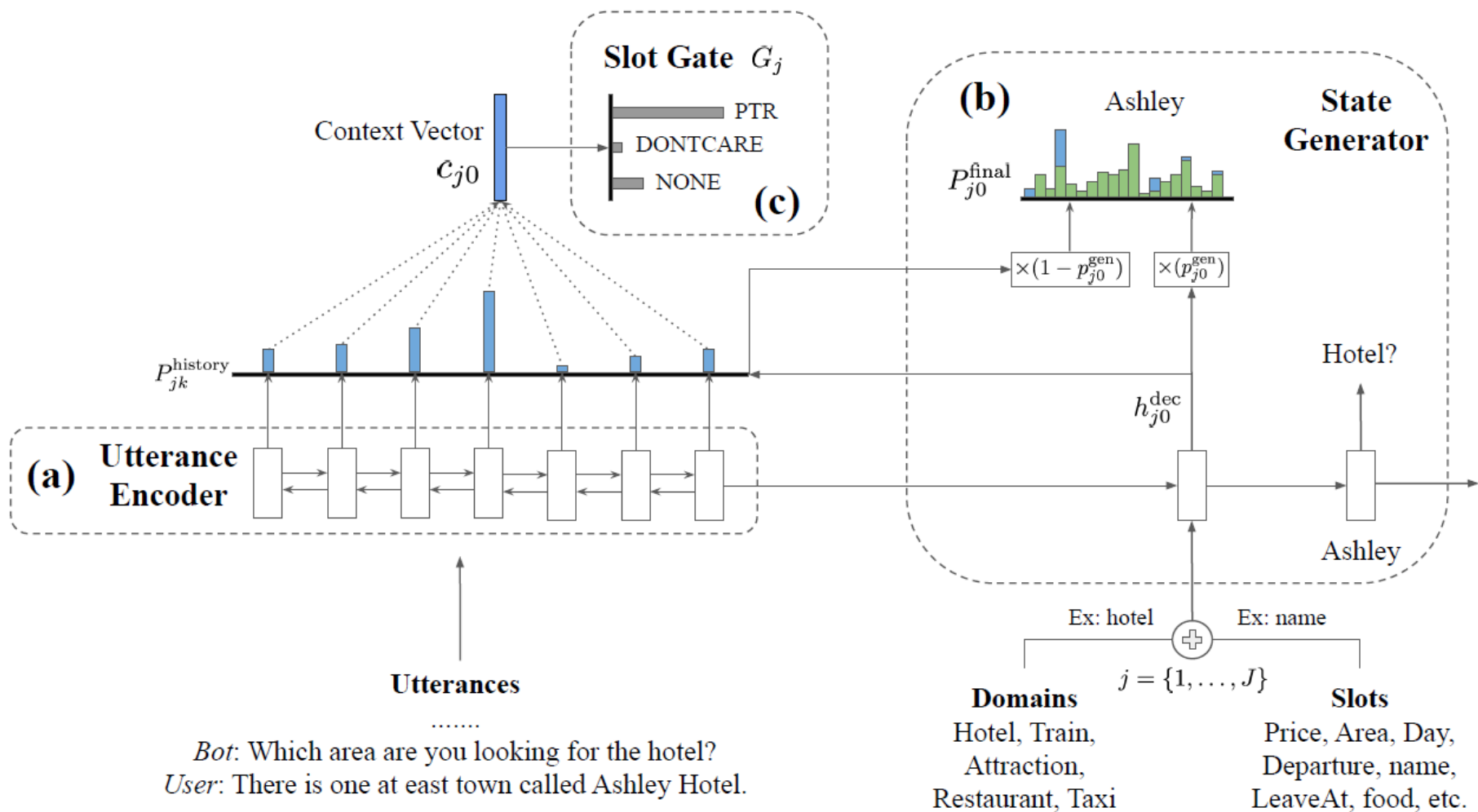|	**Hotel**	**Train**	**Attraction**	**Restaurant**
Slots	price, type, parking, stay, day, people, area, stars, internet, name	destination, departure, day, arrive by, leave at, people	area, name, type	food, price, area, name, time, day, people

Question: What is the *destination* of *train*?

Question: What is the *destination* of *taxi*?

TRADE

TRAnsferable Dialogue statE generator
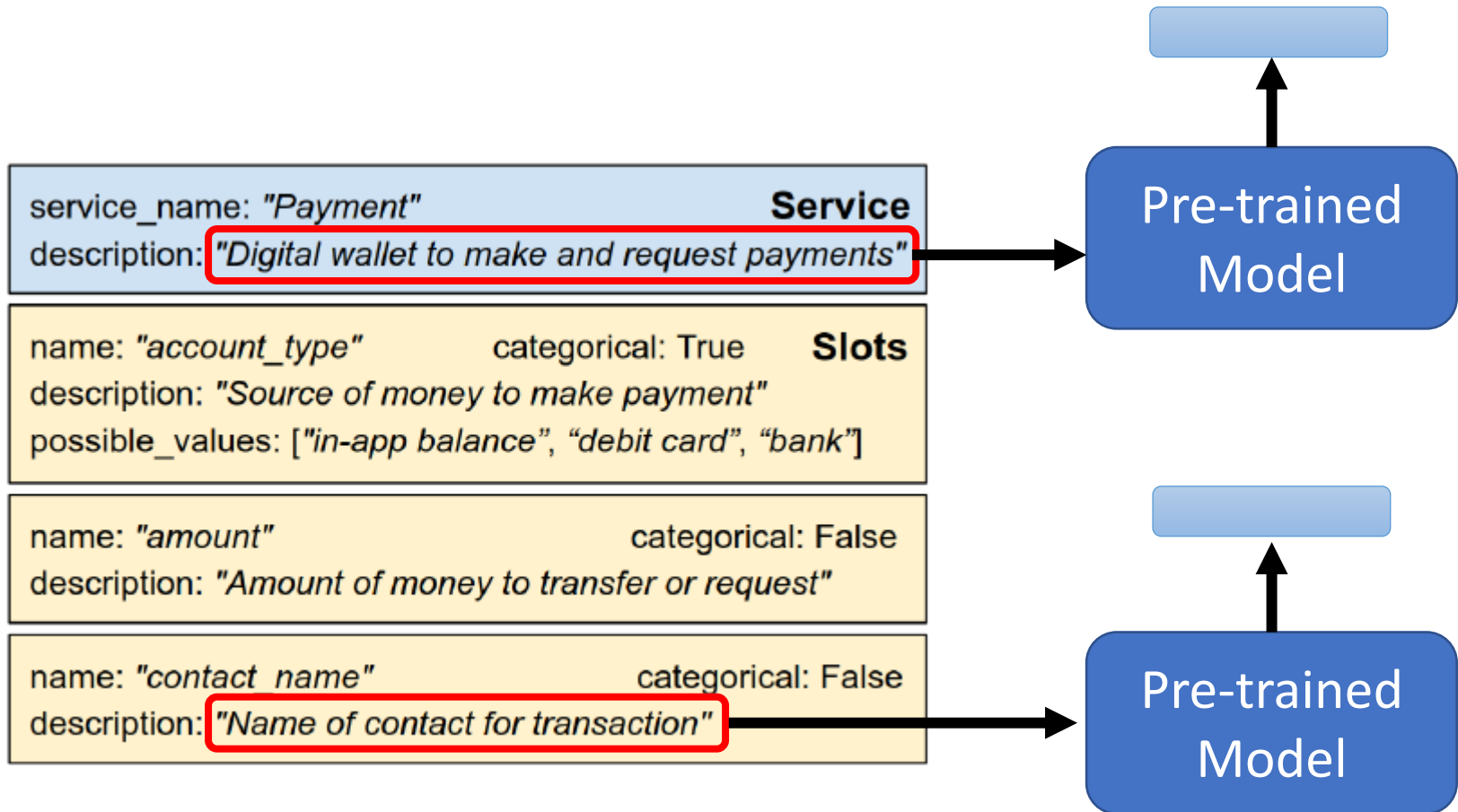
[Wu, et al., ACL'19]

(Do not need a complete question)

TRADE

- Zero Shot

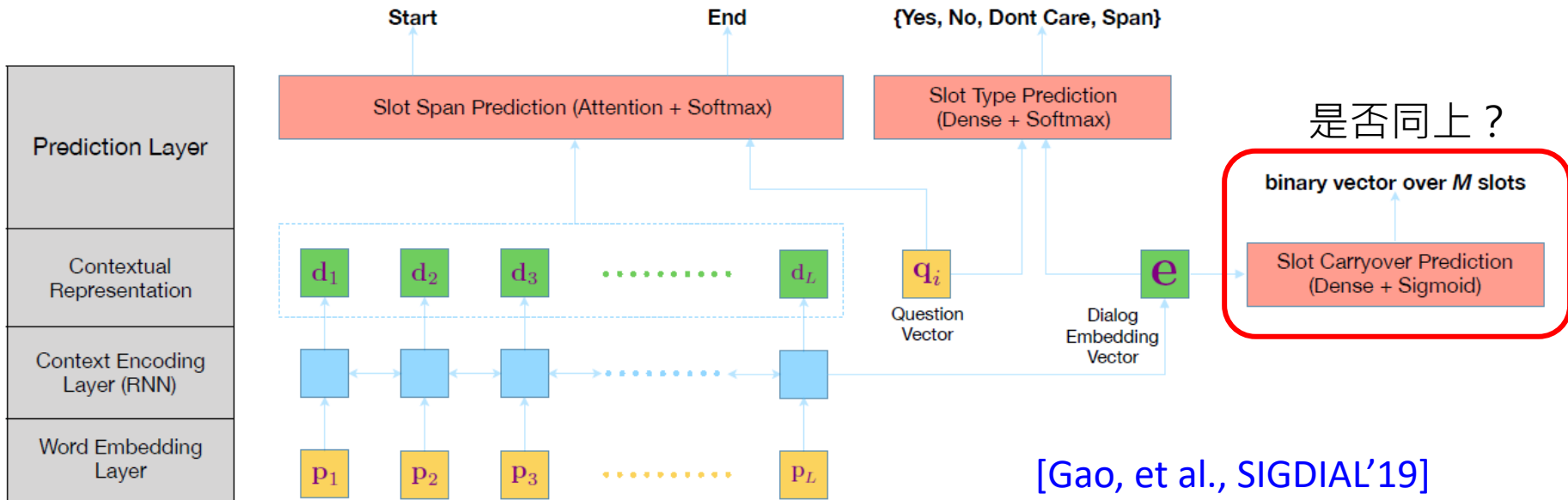|	Hotel	Train	Attraction	Restaurant	Taxi
Slots	price, type, parking, stay, day, people, area, stars, internet, name	destination, departure, day, arrive by, leave at, people	area, name, type	food, price, area, name, time, day, people	destination, departure, arrive by, leave by

	Trained Single		**Zero-Shot**	
	Joint	*Slot*	*Joint*	*Slot*
Hotel	55.52	92.66	13.70	65.32
Train	77.71	95.30	22.37	49.31
Attraction	71.64	88.97	19.87	55.53
Restaurant	65.35	93.28	11.52	53.43
Taxi	76.13	89.53	**60.58**	73.92

Embedding from Description

service_name: *"Payment"*   **Service**
description: *"Digital wallet to make and request payments"*

name: *"account_type"*   categorical: True   **Slots**
description: *"Source of money to make payment"*
possible_values: [*"in-app balance"*, *"debit card"*, *"bank"*]

name: *"amount"*   categorical: False
description: *"Amount of money to transfer or request"*

name: *"contact_name"*   categorical: False
description: *"Name of contact for transaction"*

Pre-trained Model

Pre-trained Model

[Rastogi, et al., AAAI'20]

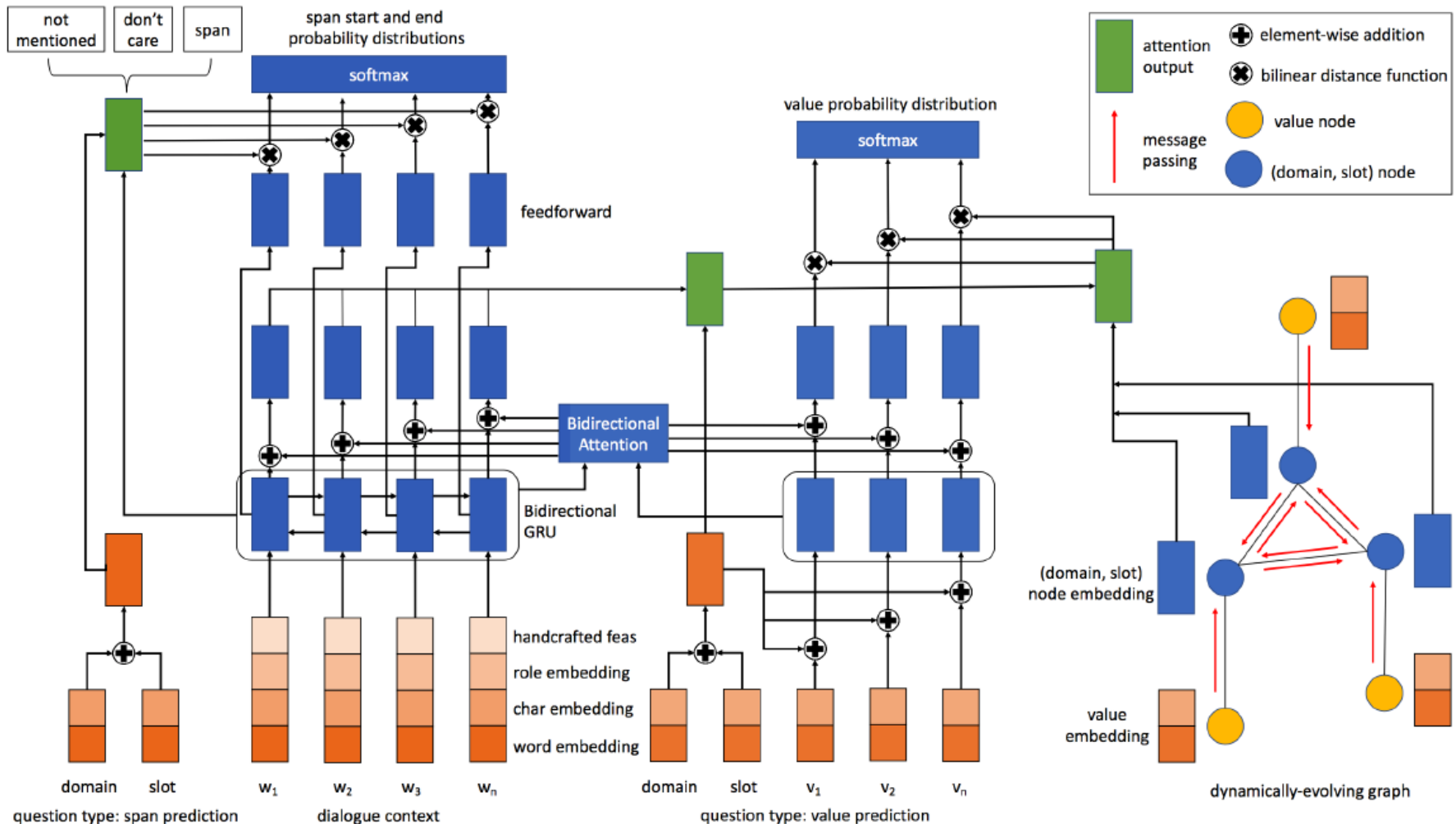# Slot Carryover Prediction

是否同上？

[Gao, et al., SIGDIAL'19]

User: I need to book a hotel in the east that has 4 stars.
area=east, stars=4 (All other slots are NULL)
Agent: I can help you with that. What is your price range?
User: That doesn't matter if it has free wifi and parking.
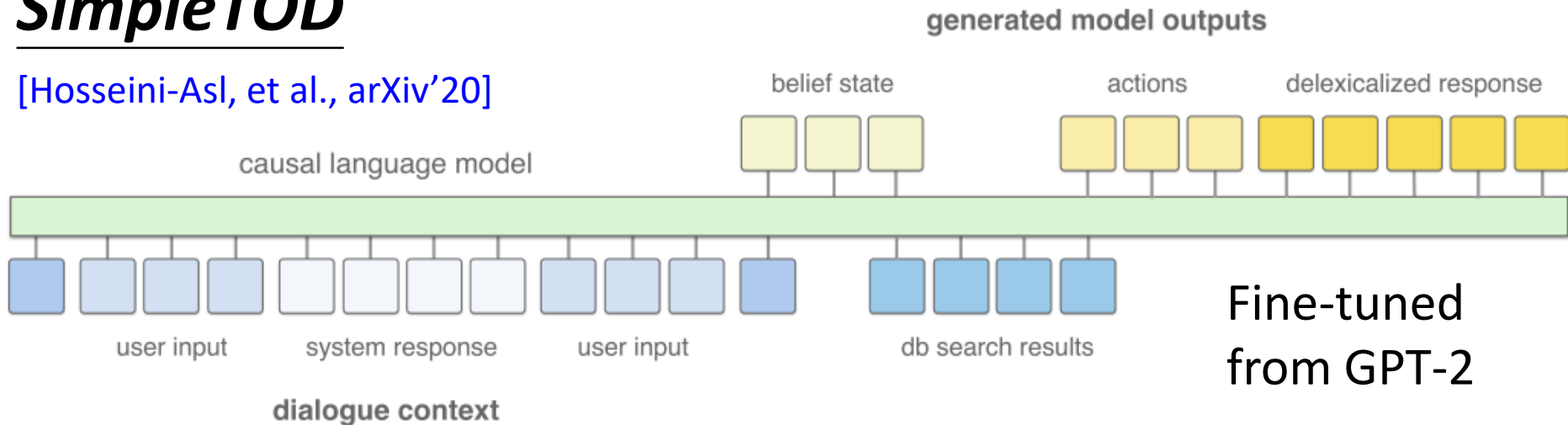parking=yes, internet=yes
price=dontcare, stars=4, area=east

DST QA

The (domain, slot) pairs are not independent.
If a user booked a restaurant, then the destination of the taxi is likely to be that restaurant.

[Zhou, et al., NeurIPS Workshop'19]

SimpleTOD

Fine-tuned from GPT-2

Context	[context] [user] *user input* [system] *system response* . . . [user] *user input* [endofcontext]
Belief State	[belief] *domain slot_name value, domain slot_name value, . . .* [endofbelief]
DB Search	[db] *#_matches, booking_status* [endofdb]
Action	[action] *domain action_type slot_name, domain action_type slot_name, . . .* [endofaction]
Response	[response] *system delexicalized response* [endofresponse]

Model	Decoder	Context Encoder	Extra Supervision	Joint Accuracy
TRADE	Generative + Classifier	Bidirectional	-	45.6
DSTQA	Classifier	Bidirectional	knowledge graph	51.17
DST-Picklist	Classifier	Bidirectional	-	53.3
SST	Generative	Bidirectional	schema graph	55.23
SimpleTOD (ours)	Generative	Unidirectional	-	**55.72**

Table 2: Evaluation of Dialogue State Tracking (DST) on MultiWOZ 2.1 using joint accuracy metric.

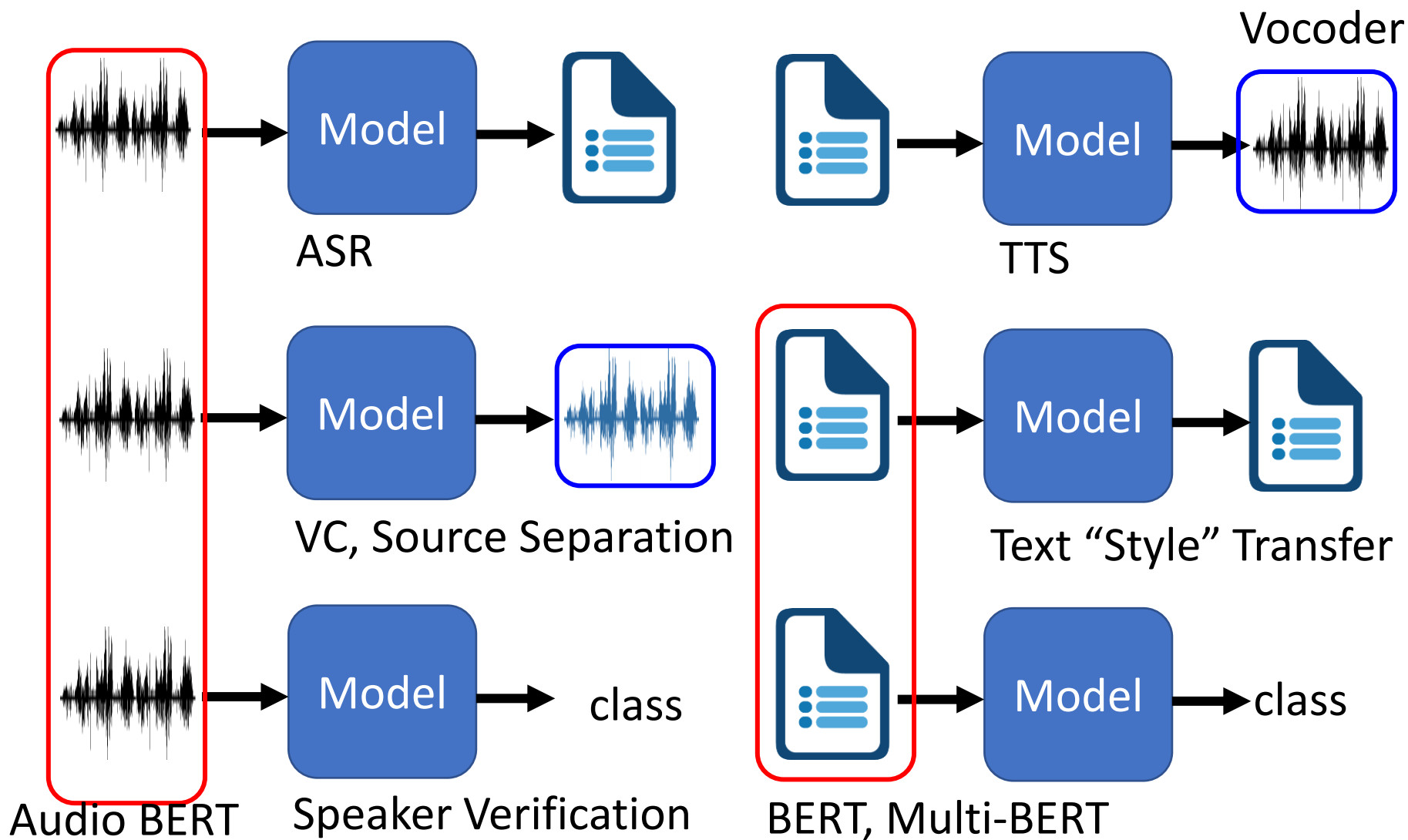|	
Context ('MUL1015', turn 10)	<|context|> <|user|> i am looking for a pool somewhere in the south of cambridge . <|system|> i am sorry , but it does not look like we have a pool that matches your criteria . <|user|> how about some type of theater ? <|system|> the junction theatre is in the south area . would you like the address ? <|user|> yes , please provide the address and postcode as well . <|system|> the junction theatre is located on clifton way at postcode cb17gx . is there anything else i can assist you with ? <|user|> thanks . yes , i also need to find a guesthouse for my stay there . <|system|> are you needing the guesthouse in the south , also ? do you have a preference in price range ? <|user|> the area should be the centre . i need it to be cheap and have free parking . <|system|> i have 2 guesthouses that fit your criteria . 1 is the alexander bed and breakfast and the other is el shaddai . <|user|> what is the star rating -s for them ? <|system|> the alexander bed and breakfast has a 4 star rating while the el shaddai has 0 stars . would you like to book a room with either of these ? <|user|> ok great , thank you . i also need to book a taxi to get back and forth . <|system|> where did you wan to leave from and go to ? <|user|> i'll be leaving the theatre at 14:45 and going to the guesthouse . <|system|> let s back up a moment . which lodgings did you prefer , so we know where the taxi will be taking you ? <|user|> let s go with the alexander bed and breakfast . <|system|> your booking is complete , a grey volkswagen will be picking you up , contact number 07753429052 . can i help with anything else ? <|user|> cool beans ! sorry i was so much trouble . thanks so much for all of your help , you need a raise dealing with customers like me ! <|endofcontext|>
Target	<|belief|> attraction area south, attraction name the junction, attraction type theatre, hotel area centre, hotel name alexander bed and breakfast, hotel parking yes, hotel pricerange cheap, hotel type guesthouse, taxi departure the junction, taxi destination alexander bed and breakfast, taxi leaveat 14:45 <|belief|> <|action|> general bye none <|endofaction|> <|response|> it was no trouble at all ! i hope you have a fantastic trip and enjoy your stay ! <|endofresponse|>
SimpleTOD	<|belief|> attraction area south, attraction name the junction, attraction type theatre, hotel area centre, hotel name alexander bed and breakfast, hotel parking yes, hotel pricerange cheap, hotel type guesthouse, taxi departure the junction, taxi destination alexander bed and breakfast, taxi leaveat 14:45 <|endofbelief|> <|action|> general bye none <|endofaction|> <|response|> thank you for using our service <|endofresponse|>

Reference

- Shuyang Gao, Abhishek Sethi, Sanchit Agarwal, Tagyoung Chung, Dilek Hakkani-Tur, Dialog State Tracking: A Neural Reading Comprehension Approach, SIGDIAL, 2019

- Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, Richard Socher, A Simple Language Model for Task-Oriented Dialogue, arXiv, 2020

- Chien-Sheng Wu, Andrea Madotto, Ehsan Hosseini-Asl, Caiming Xiong, Richard Socher, Pascale Fung, Transferable Multi-Domain State Generator for Task-Oriented Dialogue Systems, ACL, 2019

- Li Zhou, Kevin Small, Multi-domain Dialogue State Tracking as Dynamic Knowledge Graph Enhanced Question Answering, NeurIPS Workshop, 2019

- Abhinav Rastogi, Xiaoxue Zang, Srinivas Sunkara, Raghav Gupta, Pranav Khaitan, Towards Scalable Multi-domain Conversational Agents: The Schema-Guided Dialogue Dataset, AAAI, 2020

Reference

- Qi Zhu, Kaili Huang, Zheng Zhang, Xiaoyan Zhu, Minlie Huang, CrossWOZ: A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset, TACL, 2020

- Paweł Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, Milica Gašić, MultiWOZ -- A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling, EMNLP, 2018

One slide for this course

	One Sequence	**Multiple Sequences**
One Class	Sentiment Classification Stance Detection Veracity Prediction Intent Classification Dialogue Policy	NLI Search Engine Relation Extraction
Class for each Token	POS tagging Word segmentation Extractive Summarization Slotting Filling NER	
Copy from Input		Extractive QA
General Sequence	Abstractive Summarization Translation Grammar Correction NLG	General QA Chatbot State Tracker Task Oriented Dialogue
Other?	Parsing	Coreference Resolution

NAT