Applied Deep Learning

Towards Conversational AI

May 31st, 2021 http://adl.miulab.tw

National Taiwan University
國立臺灣大學

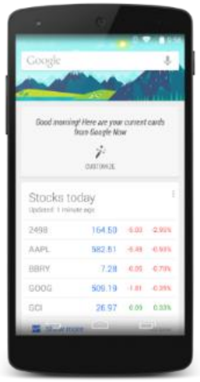
Iron Man (2008)

What can machines achieve now or in the future?

3

Language Empowering Intelligent Assistants

Apple Siri (2011)



Google Now (2012)
Google Assistant (2016)

Microsoft Cortana (2014)

Amazon Alexa/Echo (2014)

Google Home (2016)

Apple HomePod (2017)

Facebook Portal (2019)

Why Natural Language?

- Global Digital Statistics (2021 January)

Total Population
7.83B

Internet Users
4.66B (59.5%)

Unique Mobile Users
5.22B (66.6%)

Active Mobile Social Users
4.20B (53.6%)

The more **natural** and **convenient** input of devices evolves towards speech.

5

Why and When We Need?

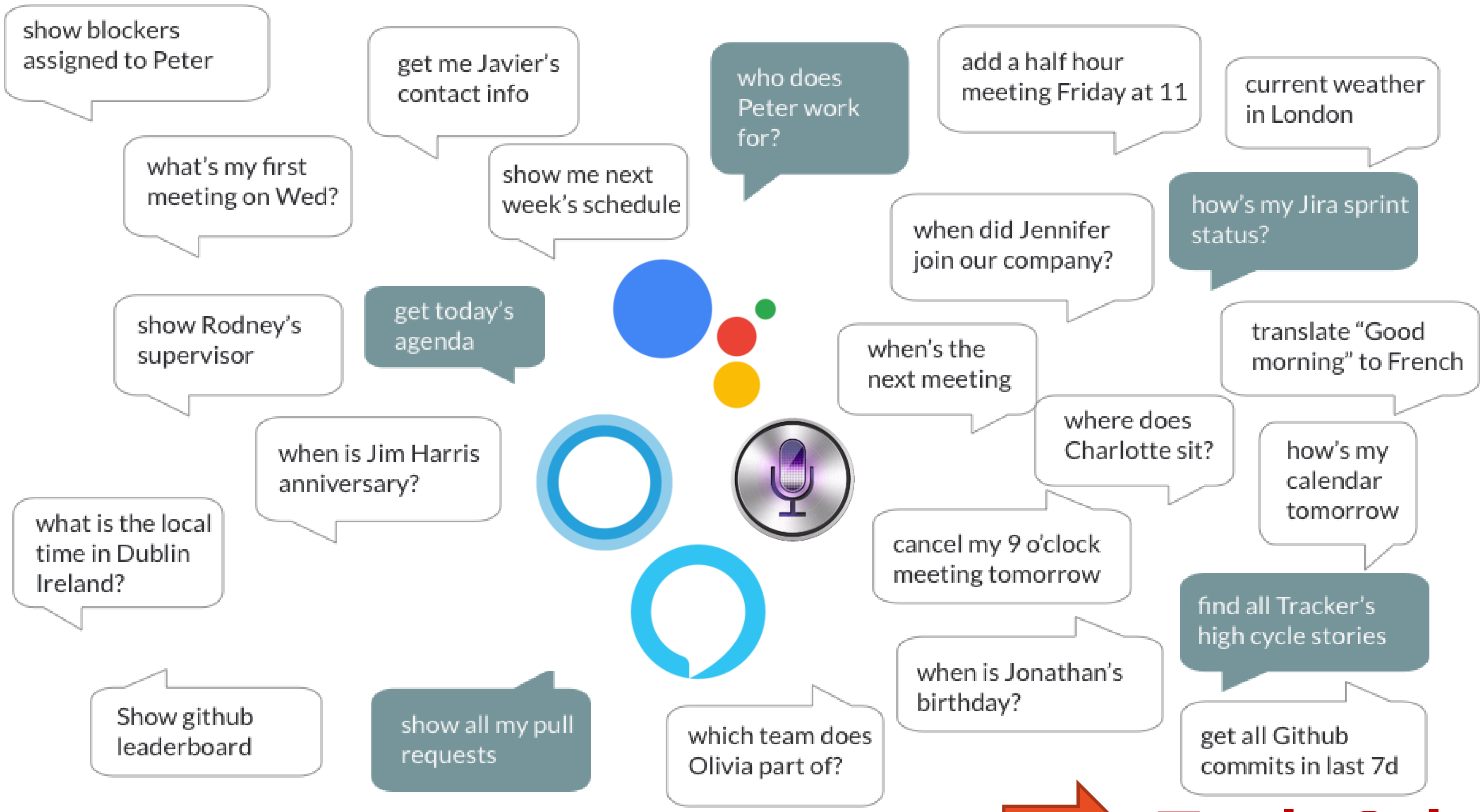
"I want to chat"	Turing Test (talk like a human)	Social Chit-Chat
"I have a question"	Information consumption	Task-Oriented Dialogues
"I need to get this done"	Task completion	
"What should I do?"	Decision support	

- *What is today's agenda?*
- *What does NLP stand for?*

- *Book me the train ticket from Kaohsiung to Taipei*
- *Reserve a table at Din Tai Fung for 5 people, 7PM tonight*
- *Schedule a meeting with Vivian at 10:00 tomorrow*

- *Is this course good to take?*

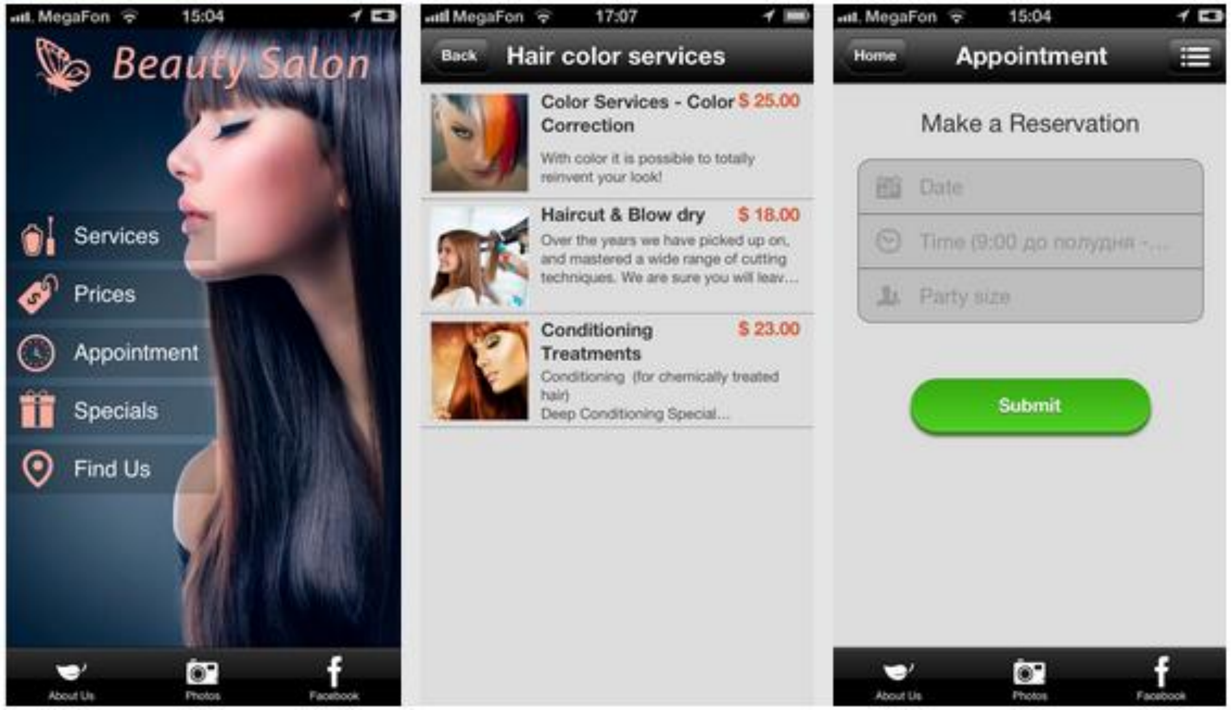
Intelligent Assistants

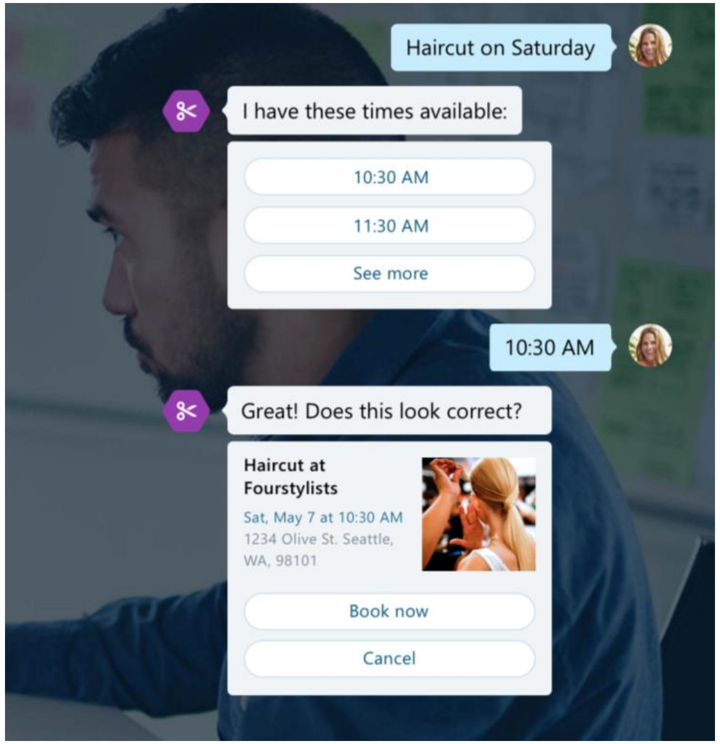


Task-Oriented

App → Bot

- A bot is responsible for a “single” domain, similar to an app





Users can initiate dialogues instead of following the GUI design

Two Branches of Conversational AI

Chit-Chat

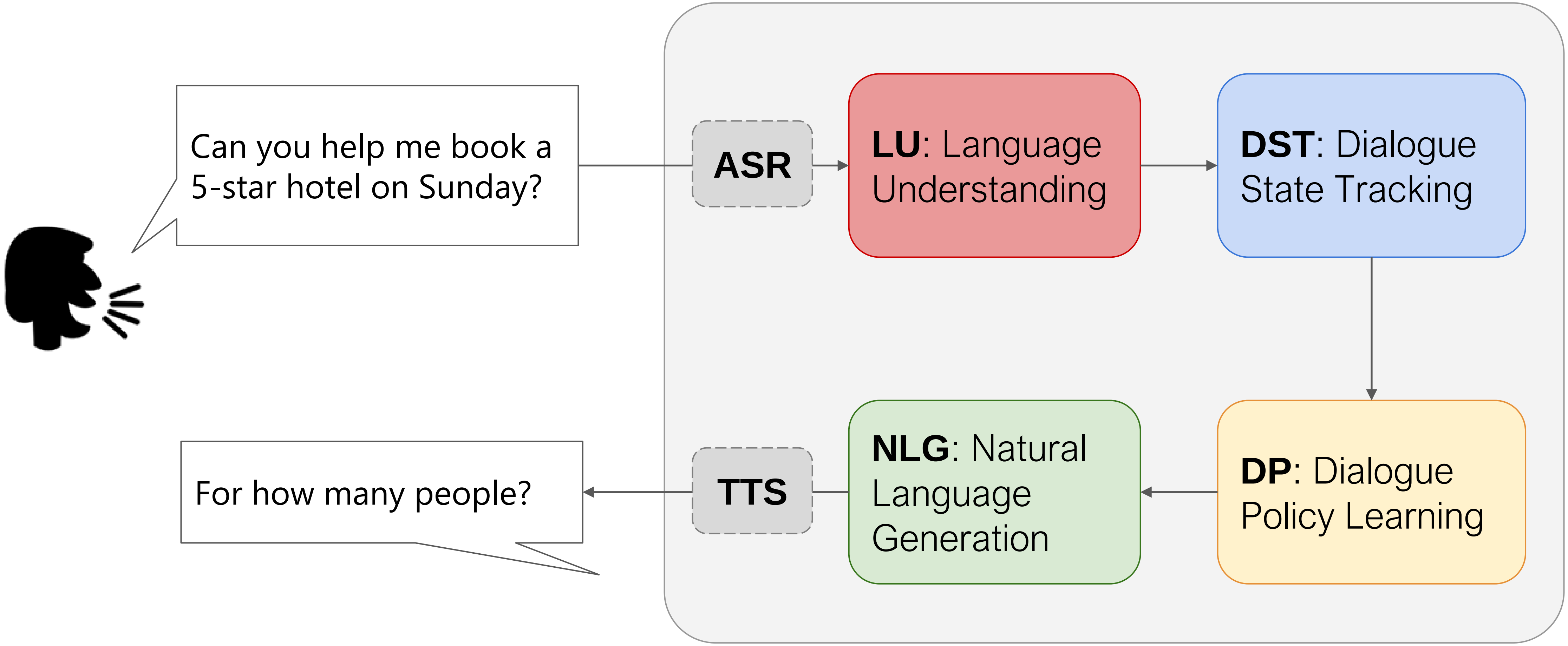
- Seq2seq models
- Seq2seq with conversation contexts
- Knowledge-grounded seq2seq models

Task-Oriented

- Single-domain, system-initiative
- Multi-domain, contextual, mixed-initiative
- End-to-end learning, massively multi-domain

Task-Oriented Dialogue Systems

Task-Oriented Dialogue Systems (Young, 2000)

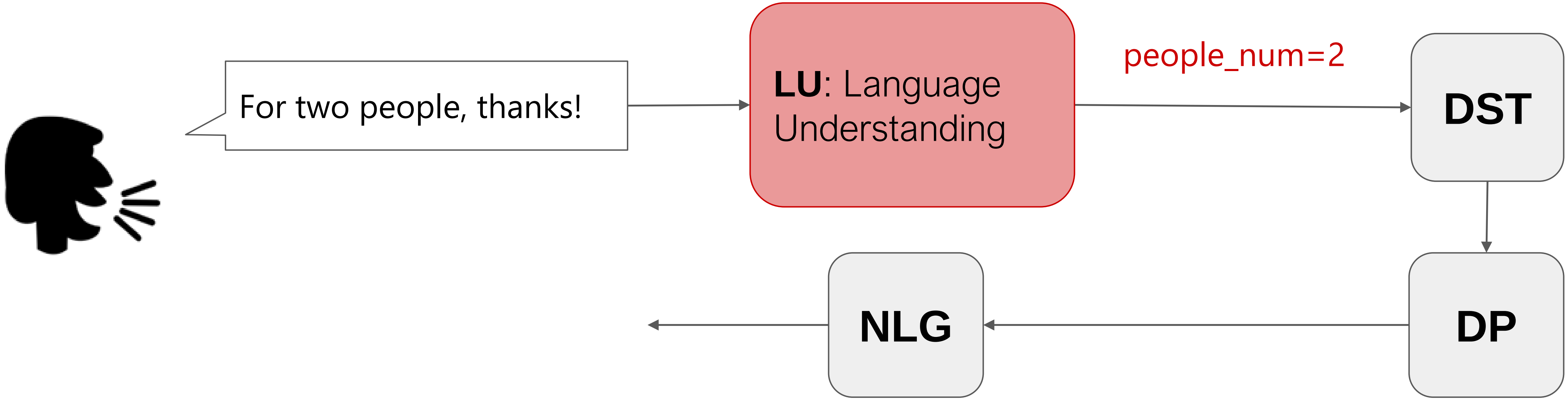


11

Language Understanding

Modular Task-Oriented Dialogue Systems

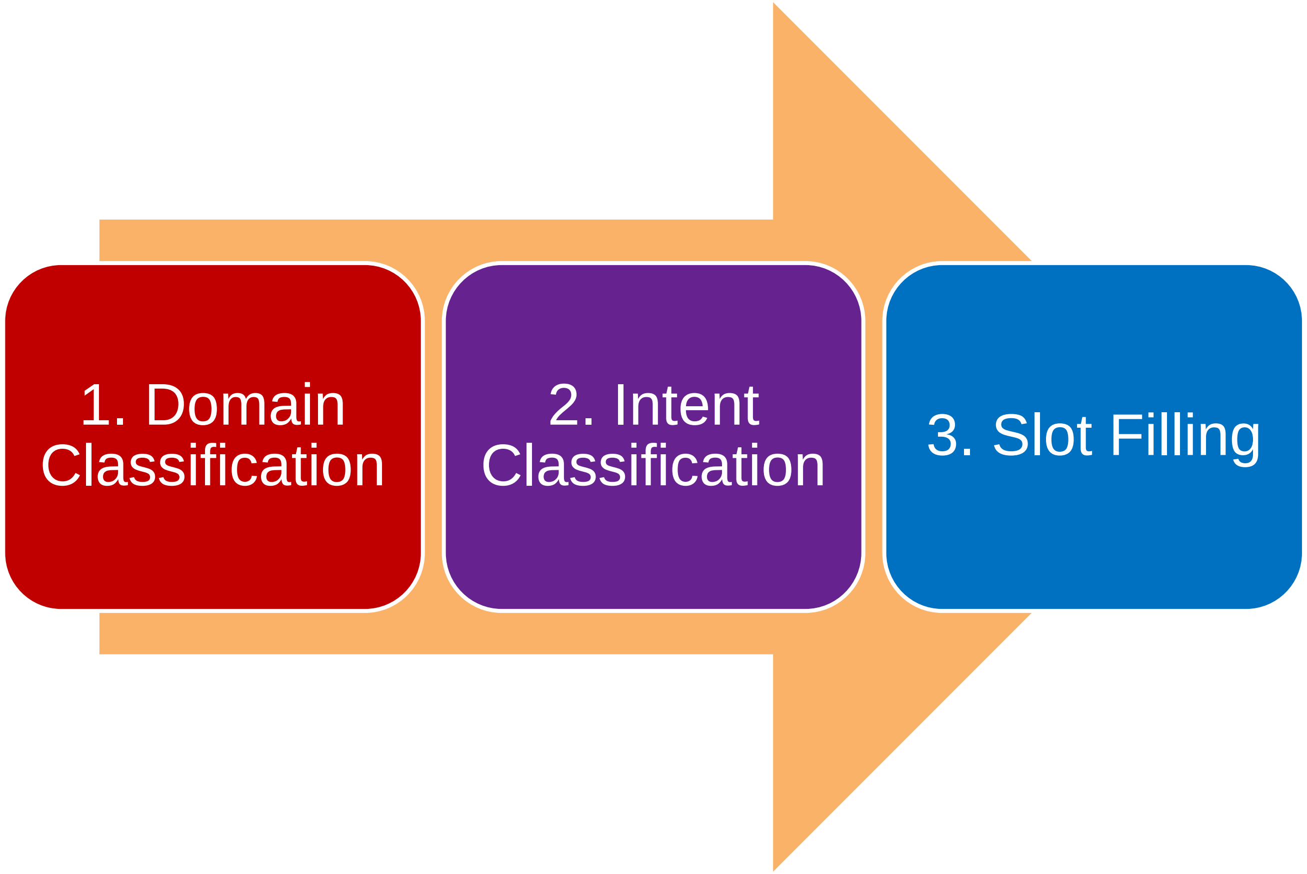
Language Understanding (LU)



- NLU is a turn-level task that maps utterances to semantics frames.
  - Input: raw user utterance
  - Output: semantic frame (e.g. speech-act, intent, slots)

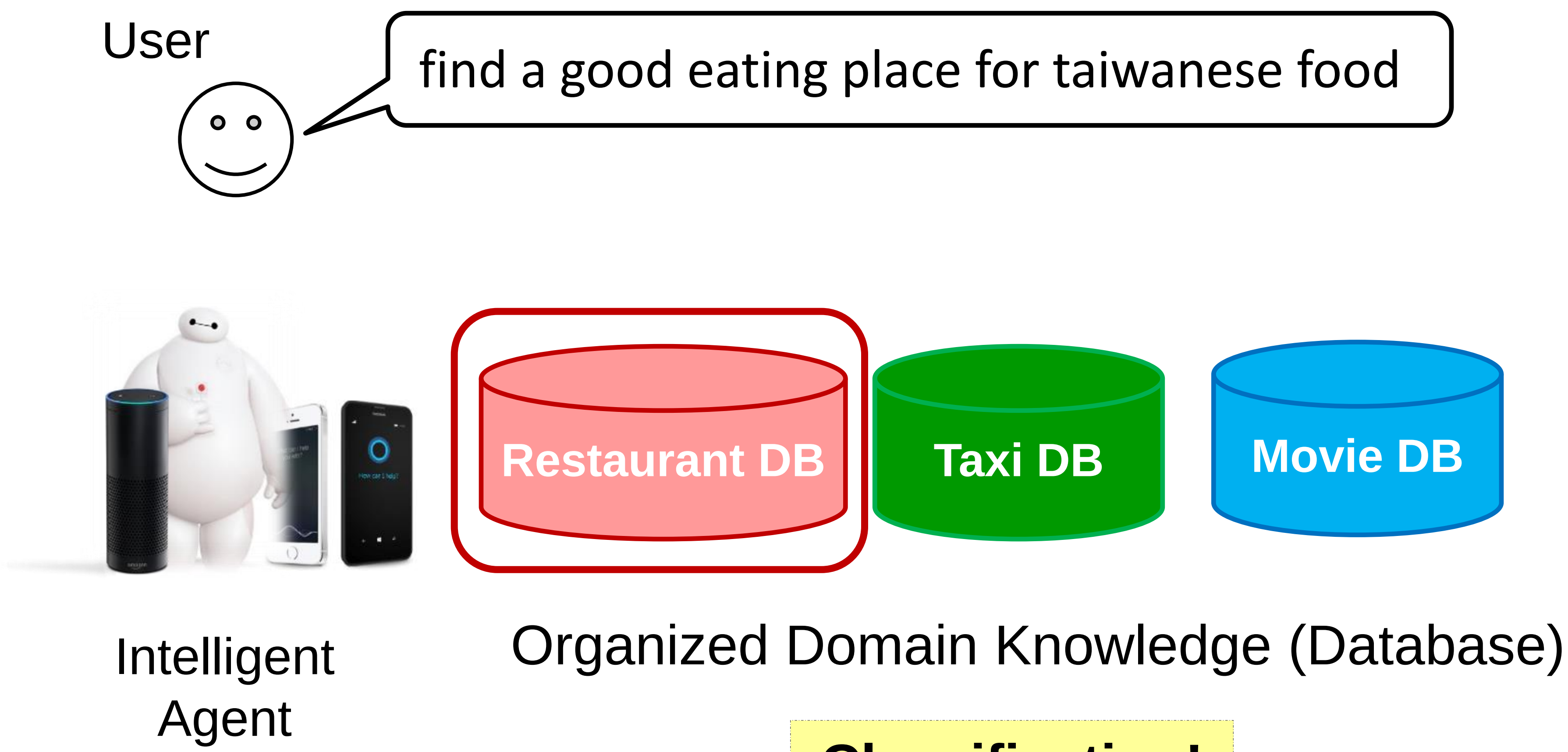
Language Understanding (LU)

- Pipelined



1. Domain Identification

Requires Predefined Domain Ontology

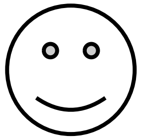


Classification!

2. Intent Detection

Requires Predefined Schema

User



find a good eating place for taiwanese food

Intelligent Agent

Restaurant DB

FIND_RESTAURANT
FIND_PRICE
FIND_TYPE
:

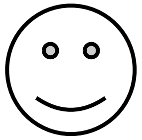
Classification!

3. Slot Filling

Requires Predefined Schema

O O B-rating O O O B-type O

User



find a good eating place for taiwanese food

Restaurant	Rating	Type
Rest 1	good	Taiwanese
Rest 2	bad	Thai
:	:	:

Restaurant DB

Intelligent Agent

FIND_RESTAURANT
rating="good"
type="taiwanese"

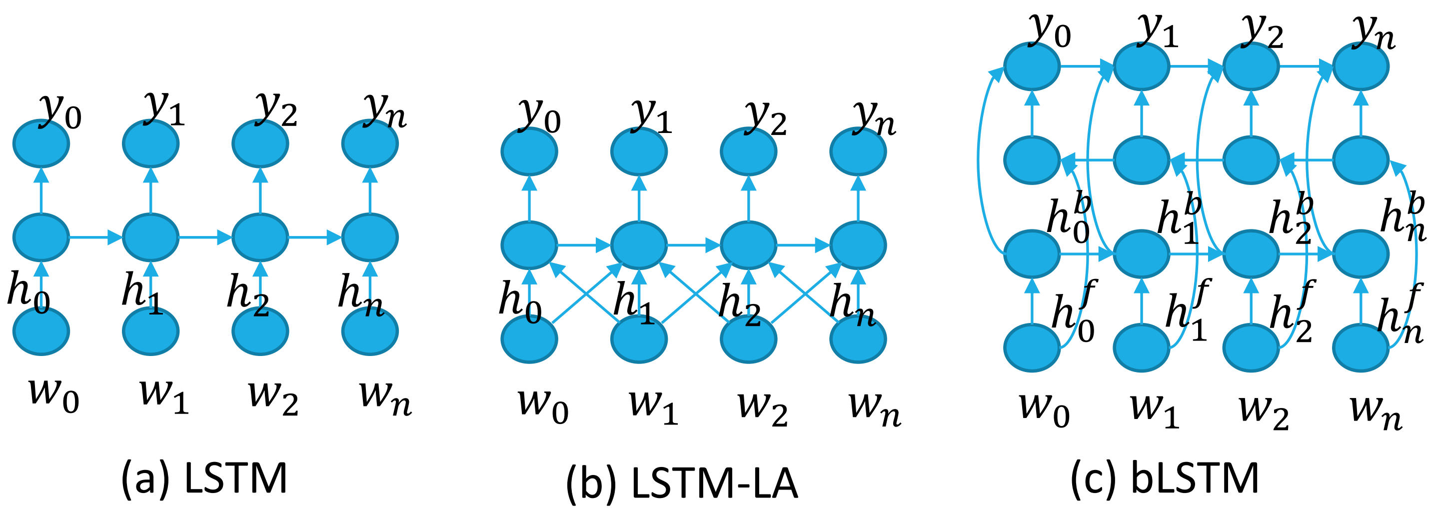
Semantic Frame

SELECT restaurant {
rest.rating="good"
rest.type="taiwanese"
}

Sequence Labeling

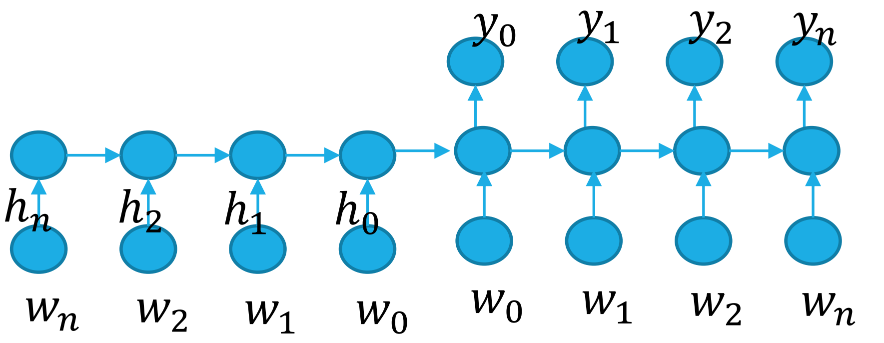
Slot Tagging (Yao et al, 2013; Mesnil et al, 2015)

- Variations:
  a. RNNs with LSTM cells
  b. Input, sliding window of n-grams
  c. Bi-directional LSTMs

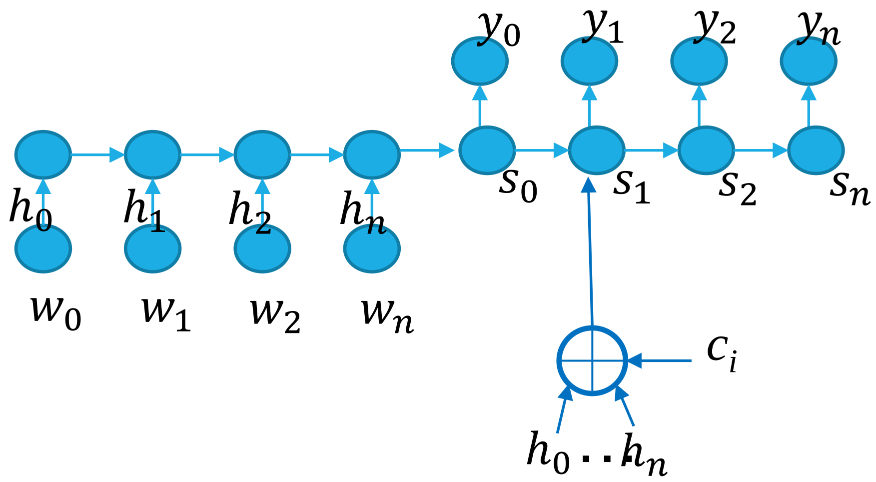


Slot Tagging (Kurata et al., 2016; Simonnet et al., 2015)

- Encoder-decoder networks
  - Leverages sentence level information

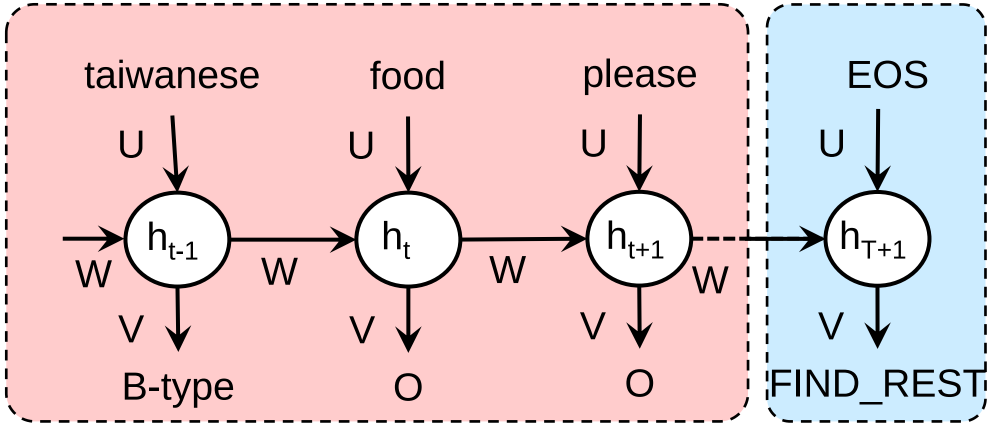


- Attention-based encoder-decoder
  - Use of attention (as in MT) in the encoder-decoder network
  - Attention is estimated using a feed-forward network with input: h_t and s_t at time t



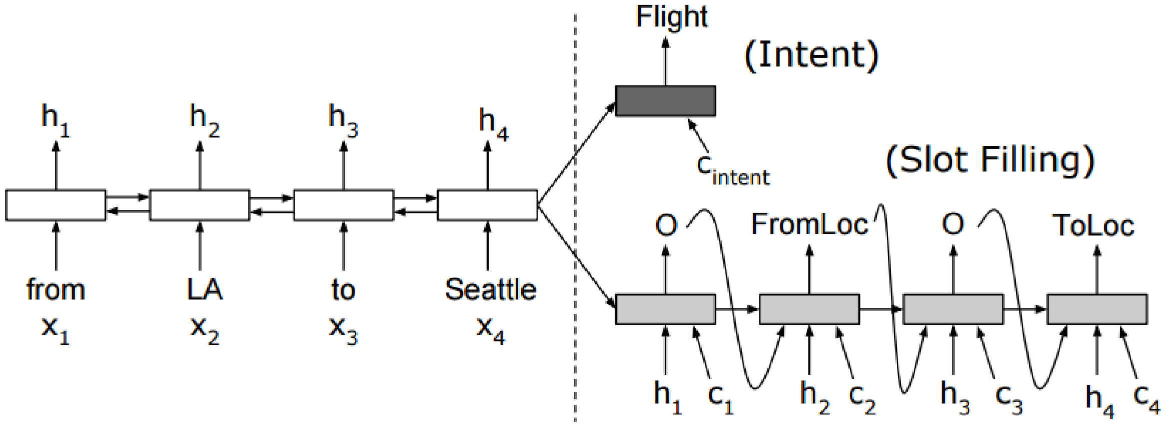
Joint Semantic Frame Parsing

- Sequence-based (Hakkani-Tur+, 2016)



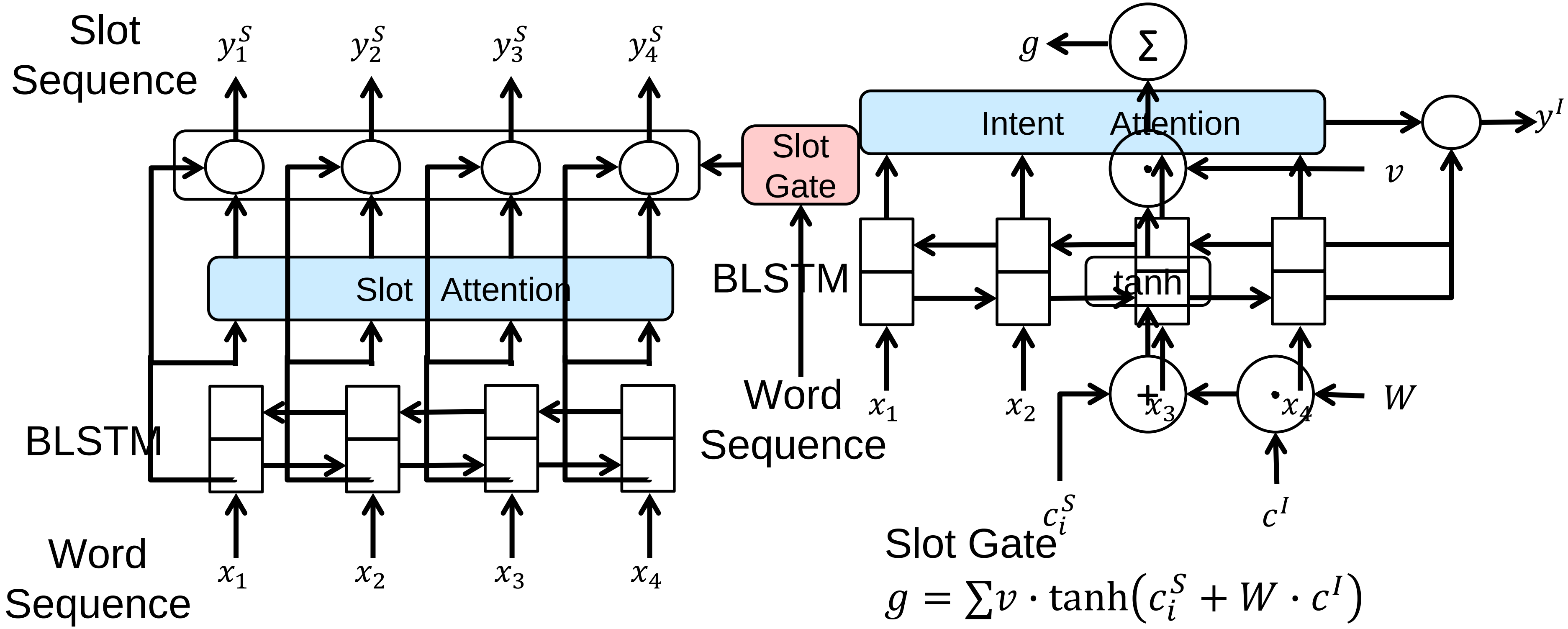
Slot Filling Intent Prediction

- Parallel-based (Liu and Lane, 2016)



	Attention Mechanism	Intent-Slot Relationship
Sequence-based (Hakkani-Tur+, '16)	X	Δ (Implicit)
Parallel-based (Liu & Lane, '16)	√	Δ (Implicit)
Slot-Gated Joint Model	√	√ (Explicit)

Slot-Gated Joint SLU (Goo+, 2018)



Slot Gate

$$g = \sum v \cdot \tanh\left(c_i^S + W \cdot c^I\right)$$

Slot Prediction

$$y_i^S = \mathrm{softmax}\left(W^S\left(h_i + \boldsymbol{g} \cdot c_i^S\right) + b^S\right)$$

$\boldsymbol{g}$ will be larger if slot and intent are better related

Contextual Language Understanding

- User utterances are highly ambiguous in isolation

Restaurant Booking

Book a table for 10 people tonight.

Which restaurant would you like to book a table for?

Cascal, for 6.

?

#people time

End-to-End Memory Networks (Sukhbaatar et al, 2015)

U: “i d like to purchase tickets to see deepwater horizon”

S: “for which theatre”

U: “angelika”

S: “you want them for angelika theatre?”

U: “yes angelika”

S: “how many tickets would you like ?”

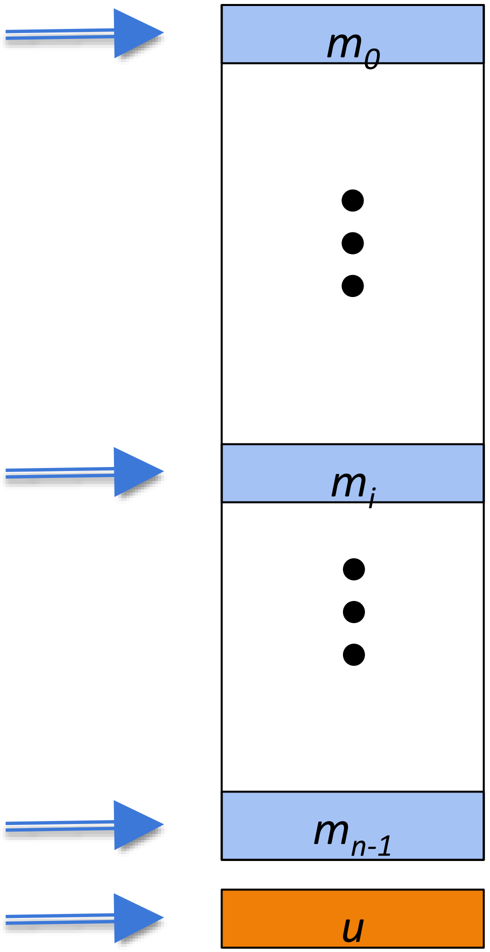
U: “3 tickets for saturday”

S: “What time would you like ?”

U: “Any time on saturday is fine”

S: “okay , there is 4:10 pm , 5:40 pm and 9:20 pm”

U: “Let’s do 5:40”



E2E MemNN for Contextual LU (Chen+, 2016)

1. Sentence Encoding

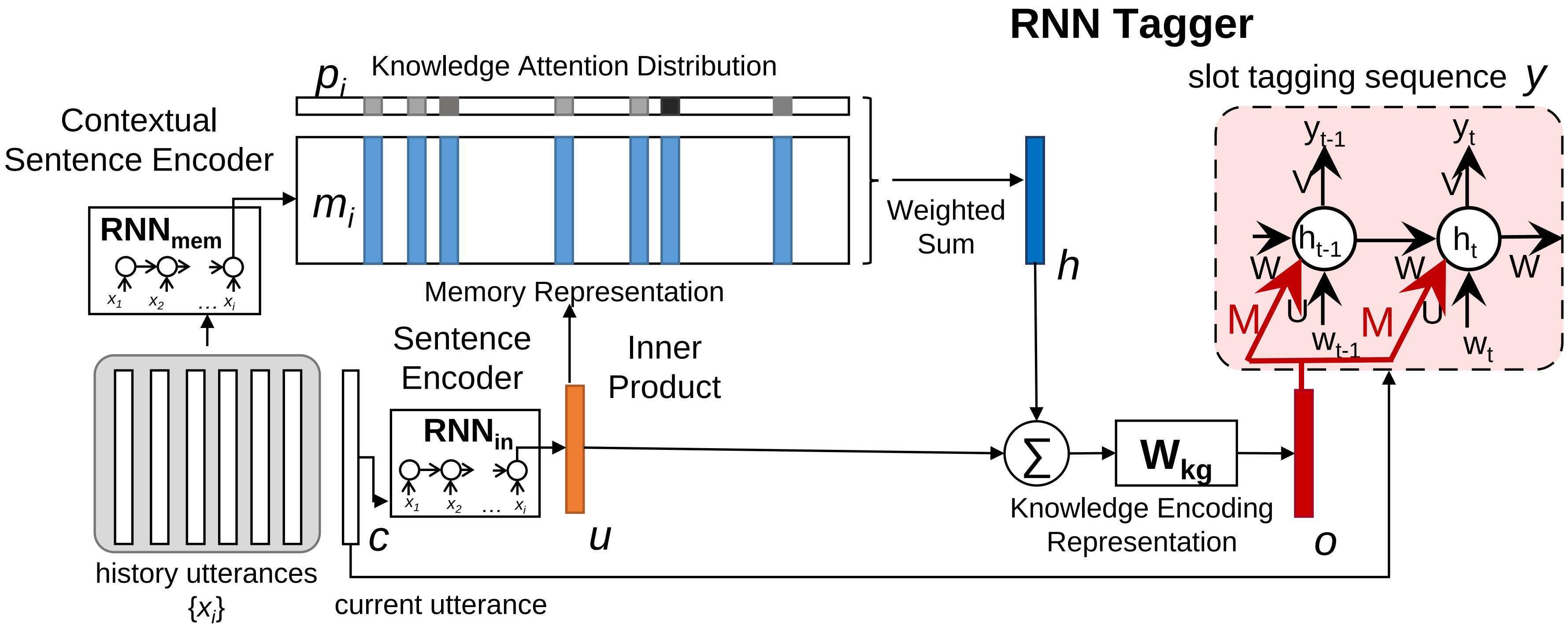
$$m_i = \mathrm{RNN}_{\mathrm{mem}}(x_i)$$
$$u = \mathrm{RNN}_{\mathrm{in}}(c)$$

2. Knowledge Attention

$$p_i = \mathrm{softmax}(u^T m_i)$$

3. Knowledge Encoding

$$h = \sum_i p_i m_i \quad o = W_{\mathrm{kg}}(h + u)$$



Idea: additionally incorporating contextual knowledge during slot tagging

E2E MemNN for Contextual LU (Chen+, 2016)

U: *"i d like to purchase tickets to see deepwater horizon"* → 0.69

S: *"for which theatre"*

U: *"angelika"*

S: *"you want them for angelika theatre?"*

U: *"yes angelika"*

S: *"how many tickets would you like ?"* → 0.13

U: *"3 tickets for saturday"*

S: *"What time would you like ?"*

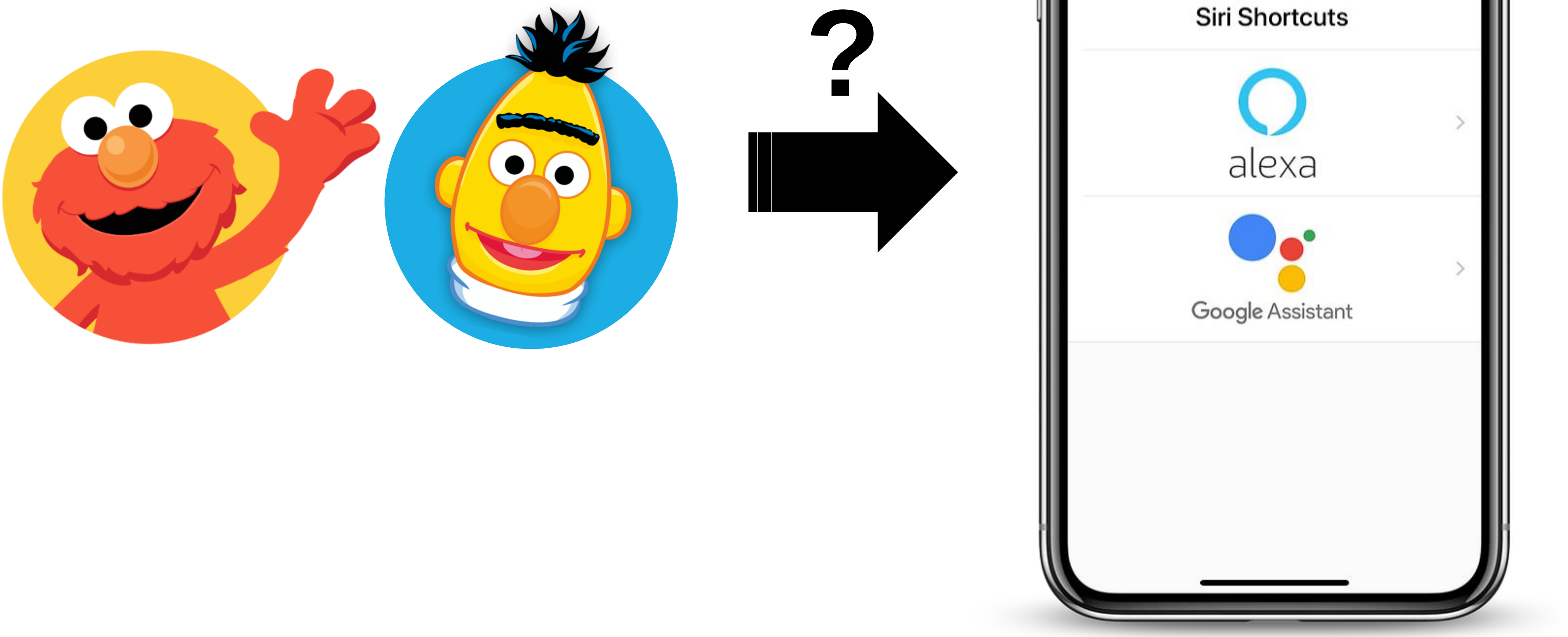
U: *"Any time on saturday is fine"*

S: *"okay , there is 4:10 pm , 5:40 pm and 9:20 pm"* → 0.16

U: "Let's do 5:40"

Recent Advances in NLP

- Contextual Embeddings (ELMo & BERT)
  - Boost many understanding performance with pre-trained language models

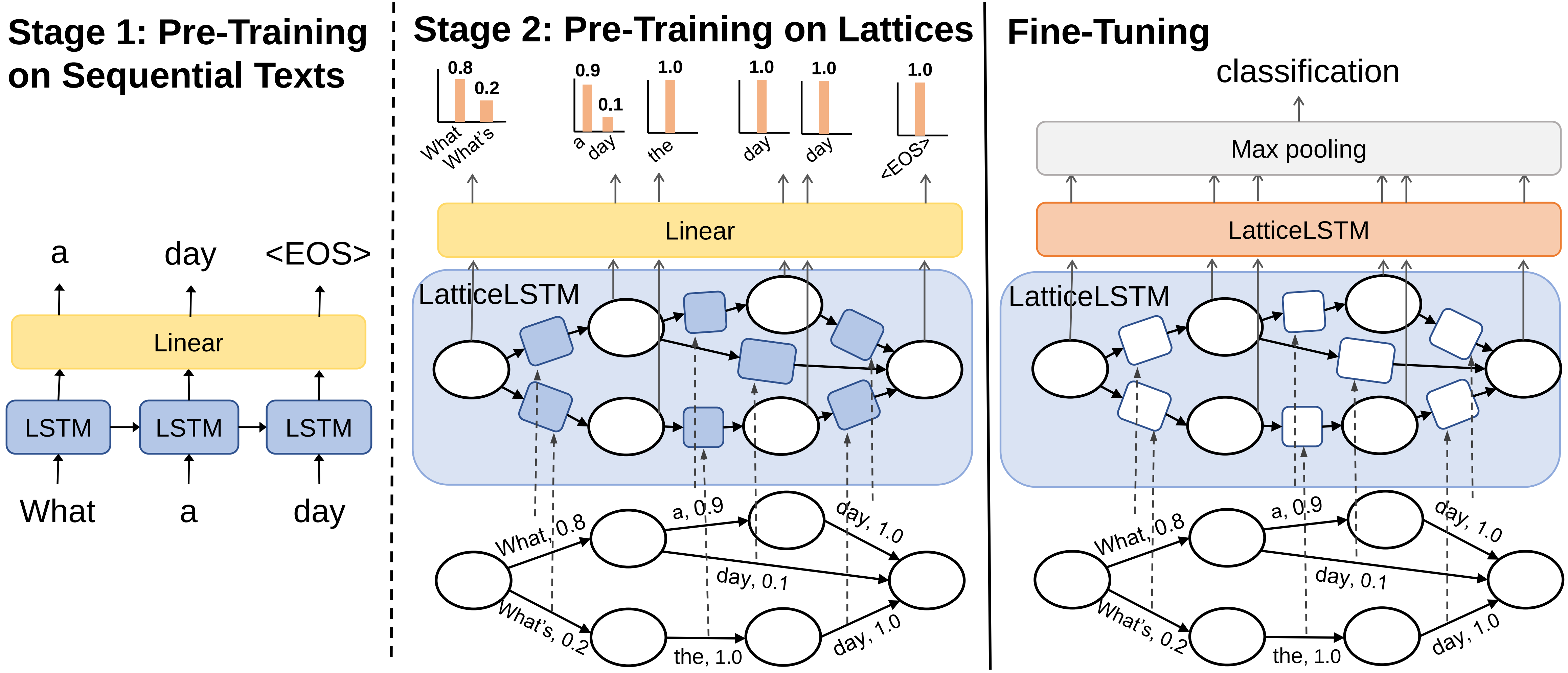



Call me ASAP

~~台灣行政機關廢言的本部案例~~

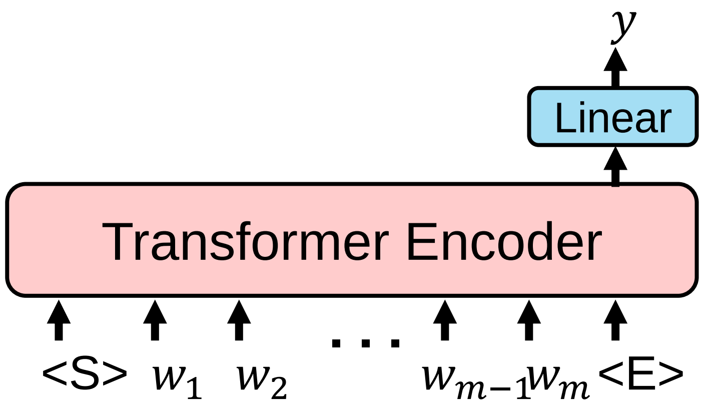
台灣有新增幾個肺炎的本土案例?

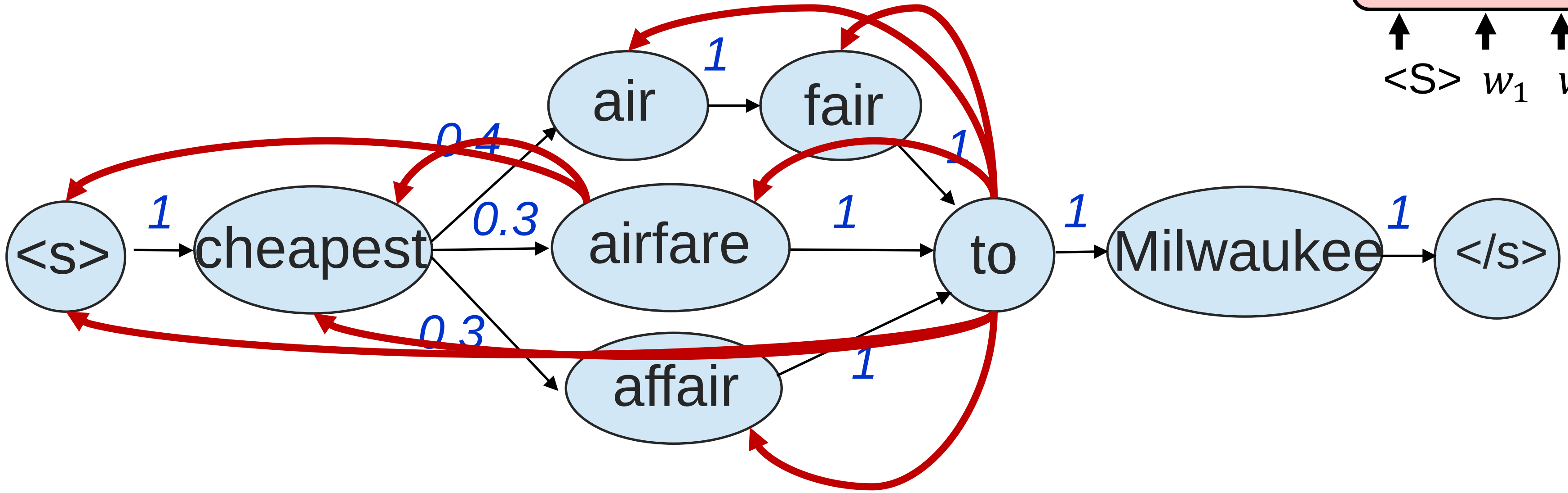
Robustness – Adapting to ASR (Huang & Chen, 2019)



Robustness – Adapting to ASR (Huang & Chen, 2019)

- Idea: lattices may include correct words
- Goal: feed lattices into Transformer





$$A(Q,K,V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}} + M\right)V$$

SLU performance is improved by leveraging the lattices without increasing training/inference time

Chao-Wei Huang and Yun-Nung Chen, "Adapting Pretrained Transformer to Lattices for Spoken Language Understanding," in *Proceedings of 2019 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 2019.

Robustness – Adapting to ASR (Huang & Chen, 2019)

- Confusion-Aware Fine-Tuning
  - Supervised

Acoustic Confusion $C = \{w_3^{x_{trs}}, w_2^{x_{asr}}\}$

x_{trs} : Show me the <u>fares</u> from Dallas to Boston

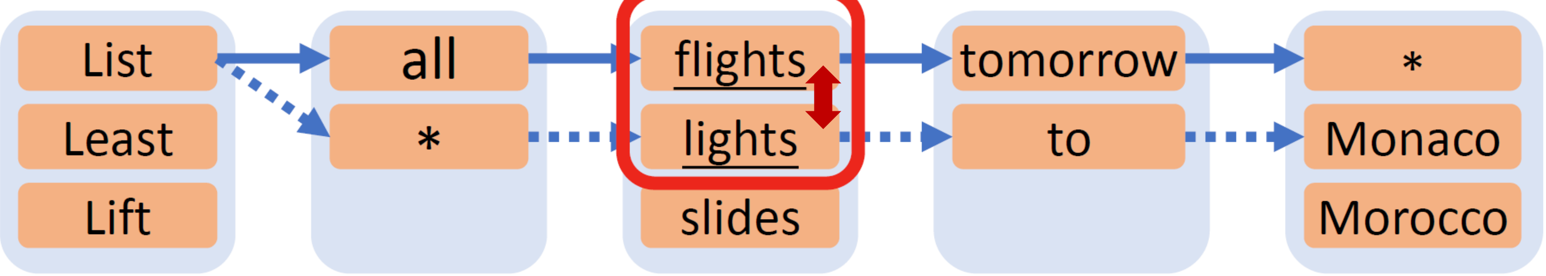
x_{asr} : Show me * <u>affairs</u> from Dallas to Boston

  - Unsupervised

Acoustic Confusion

Top hypothesis $\mathbf{x_1}$

Alternative hypothesis $\mathbf{x_2}$



The contextual embeddings of the recognized texts would be similar to the ground truth one.

Chao-Wei Huang and Yun-Nung Chen, "Learning ASR-Robust Contextualized Embeddings for Spoken Language Understanding," in *ICASSP*, 2019.

Scalability – Multilingual LU (Upadhyay+, 2018)

- Source language: English (full annotations)
- Target language: Hindi (limited annotations)

RT: round trip, FC: from city, TC: to city, DDN: departure day name

Utt:	find	a	one	way	flight	from	boston	to	atlanta	on	wednesday
Slots:	O	O	B-RT	I-RT	O	O	B-FC	O	B-TC	O	B-DDN

(a) English Utterance

Utt:	बुधवार	को	बोसटन	से	अटलांटा	तक	जाने	वाली	एकतरफ़ा	उड़ाने	खोजें
Slots:	B-DDN	O	B-FC	O	B-TC	O	O	O	B-RT	O	O

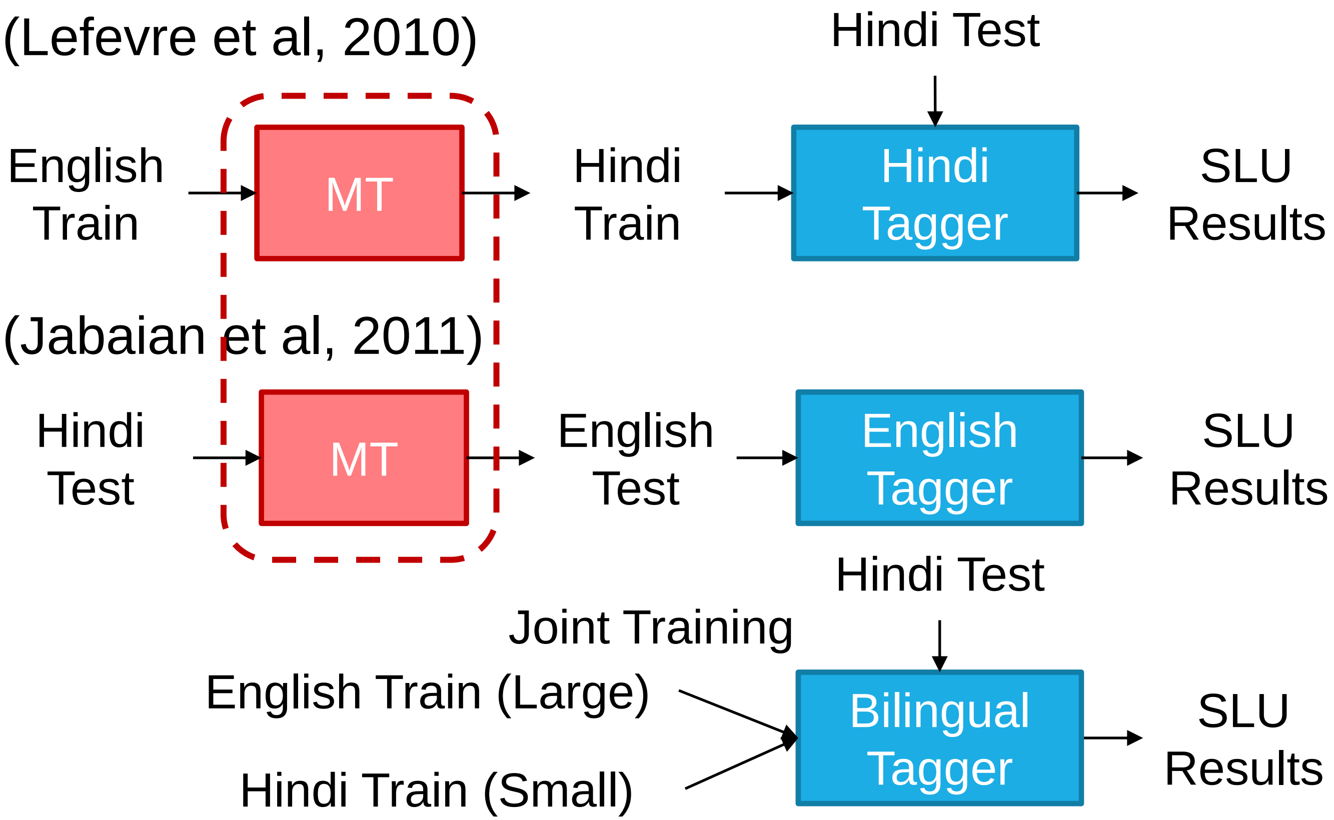
(b) Hindi Utterance

Scalability – Multilingual LU (Upadhyay+, 2018)

Train on Target (Lefevre et al, 2010)

Test on Source (Jabaian et al, 2011)

Joint Training



MT system is not required and both languages can be processed by a single model

LU Evaluation

Utterance:	For	2	people	thanks
	↓	↓	↓	↓
Slot:	O	B-people	O	O
Domain:	Hotel			
Intent:	Hotel_Book			

⇒ Slot-F1

⇒ Acc

⇒ Frame Accuracy

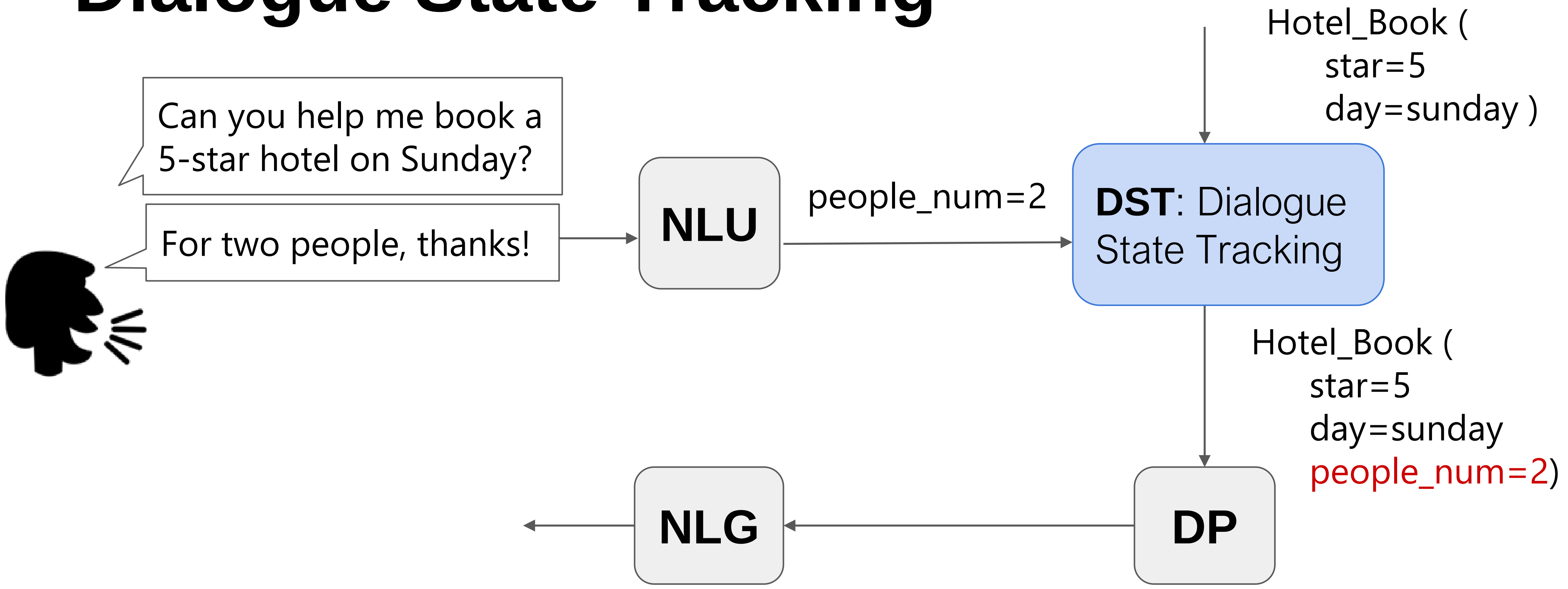
- Metrics
  - Sub-sentence-level: domain/intent accuracy, slot F1
  - Sentence-level: whole frame accuracy

34

Dialogue State Tracking

Modular Task-Oriented Dialogue Systems

Dialogue State Tracking



- DST is a dialogue-level task that maps partial dialogues into dialogue states.
  - Input: a dialogue / a turn with its previous state
  - Output: dialogue state (e.g. slot-value pairs)

Dialogue State Tracking

Hello, how may I help you?

I'm looking for a Thai restaurant. — request (restaurant; foodtype=Thai)

What part of town do you have in mind?

Something in the centre. — inform (area=centre)

Bangkok city is a nice place, it is in the centre of town and it serves Thai food.

What's the address? — request (address)

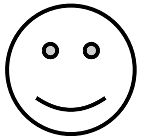
Bangkok city is a nice place, their address is 24 Green street.

Thank you, bye. — bye ()

Dialogue State Tracking

Requires Hand-Crafted States

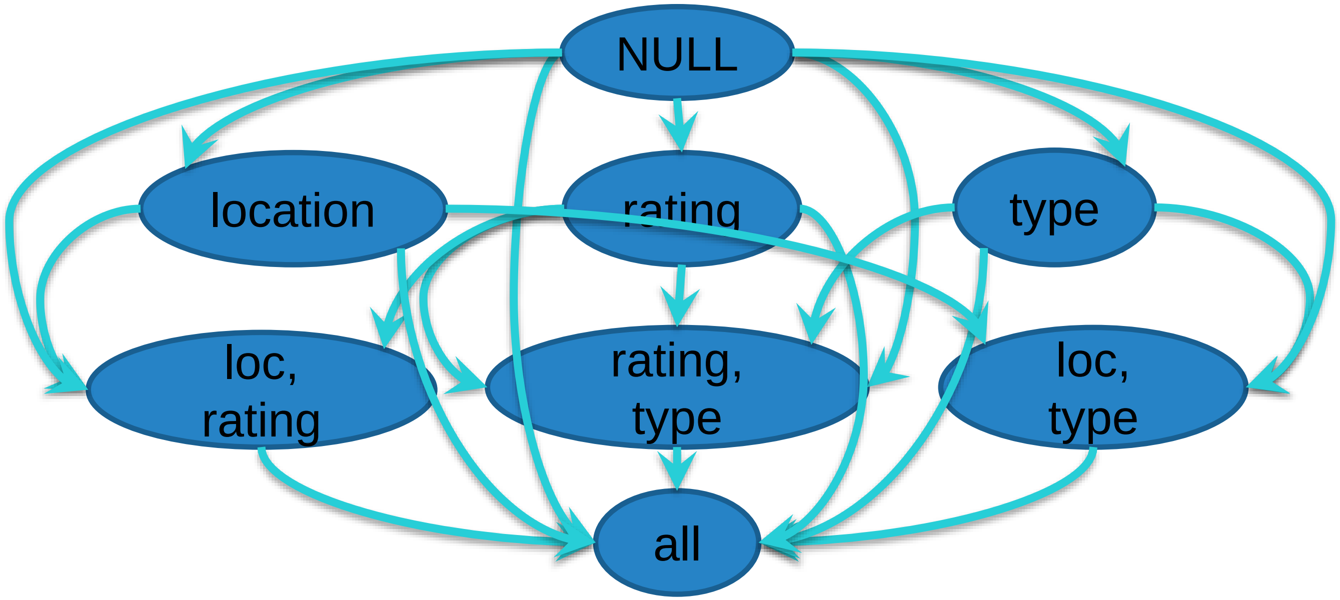
User



find a good eating place for taiwanese food

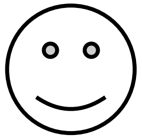
i want it near to my office

Intelligent Agent



Dialogue State Tracking
Requires Hand-Crafted States

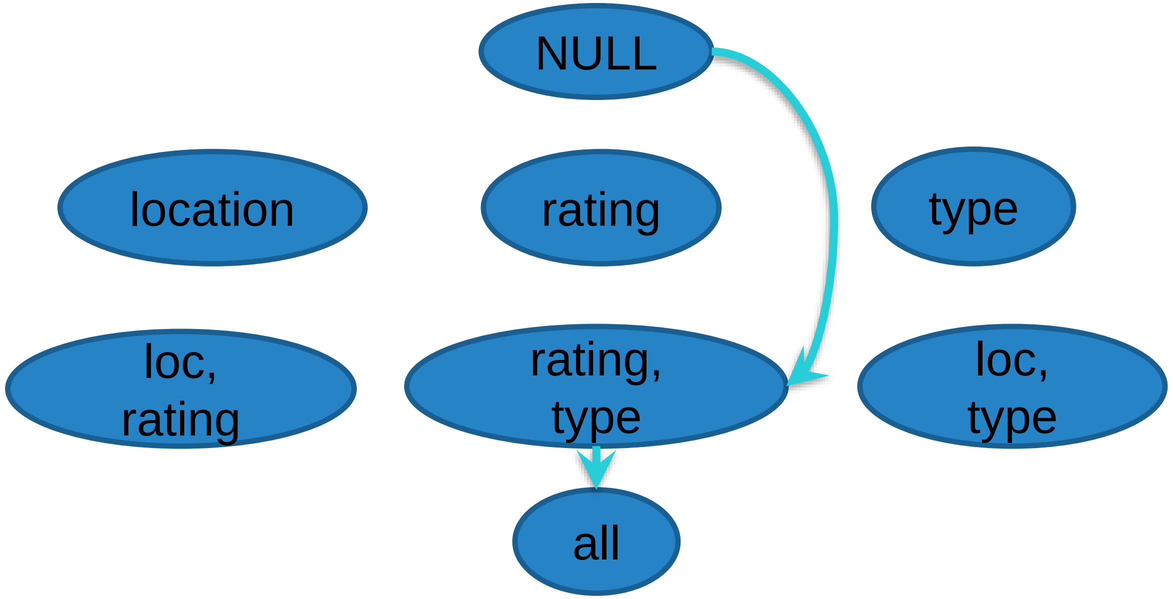
User



find a good eating place for taiwanese food

i want it near to my office

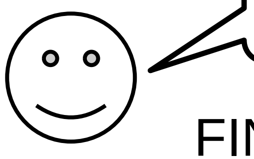
Intelligent
Agent



Dialogue State Tracking

Handling Errors and Confidence

User



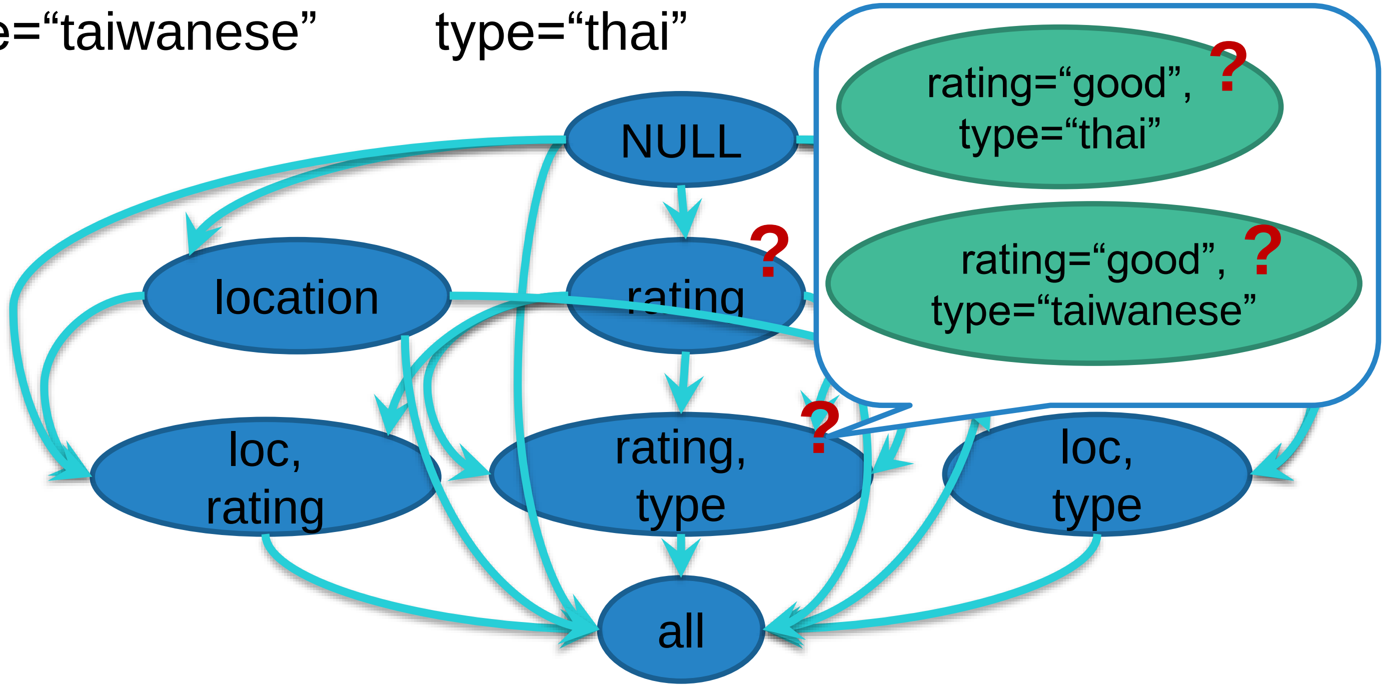
find a good eating place for taixxxx food

FIND_RESTAURANT
rating="good"
type="taiwanese"

FIND_RESTAURANT
rating="good"
type="thai"

FIND_RESTAURANT
rating="good"

Intelligent Agent



DST Problem Formulation

- The DST dataset consists of
  - *Goal*: for each informable slot
    - e.g. price=cheap
  - *Requested*: slots by the user
    - e.g. moviename
  - *Method*: search method for entities
    - e.g. by constraints, by name
- The dialogue state is
  - the distribution over possible slot-value pairs for goals
  - the distribution over possible requested slots
  - the distribution over possible methods

Dialogue State Tracking (DST)

- Maintain a probabilistic distribution instead of a 1-best prediction for better robustness to SLU errors or ambiguous input

Slot	Value
# people	5 (0.5)
time	5 (0.5)

Slot	Value
# people	3 (0.8)
time	5 (0.8)

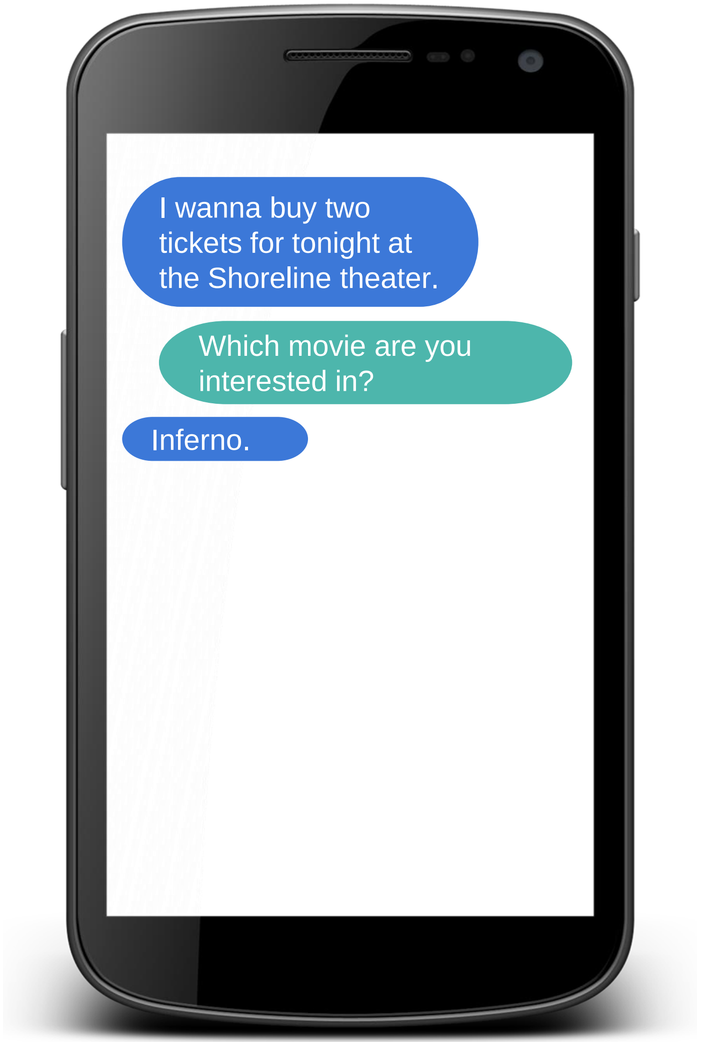
Multi-Domain Dialogue State Tracking

- A full representation of the system's belief of the user's goal at any point during the dialogue
- Used for making API calls

Movies

Date	11/15/17			
Time	6 pm	7 pm	8 pm	9 pm
#People	2			
Theater	Century 16 Shoreline			
Movie	Inferno			

Less Likely ... More Likely



Multi-Domain Dialogue State Tracking

- A full representation of the system's belief of the user's goal at any point during the dialogue
- Used for making API calls

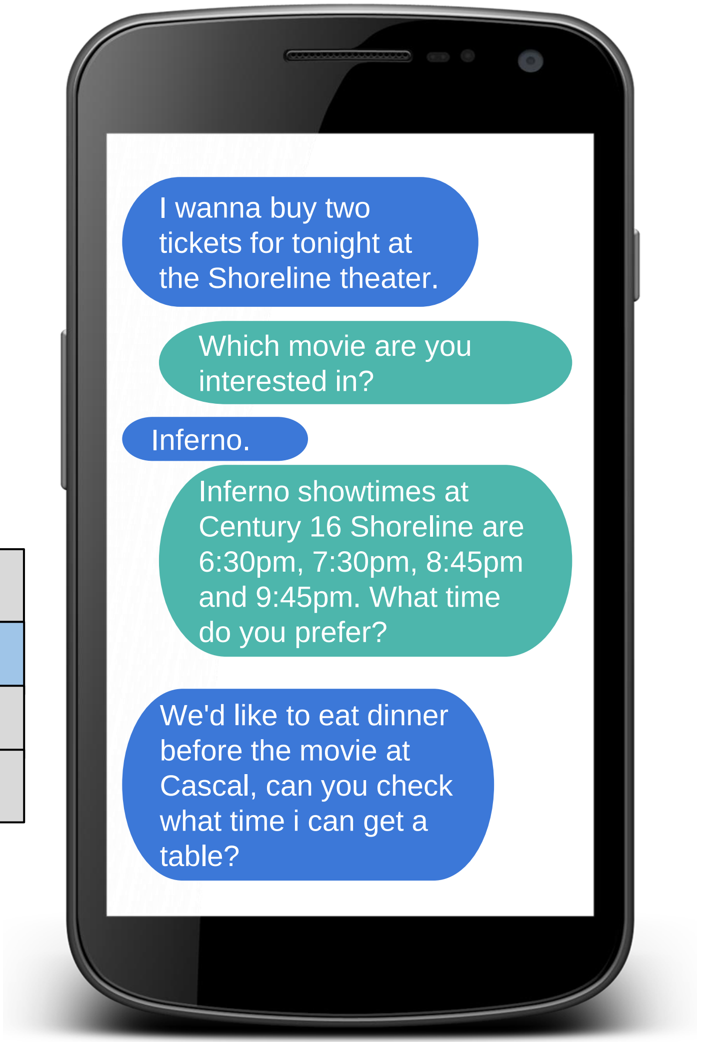
Movies

Date	11/15/17			
Time	6:30 pm	7:30 pm	8:45 pm	9:45 pm
#People	2			
Theater	Century 16 Shoreline			
Movie	Inferno			

Restaurants

Date	11/15/17		
Time	6:00 pm	6:30 pm	7:00 pm
Restaurant	Cascal		
#People	2		

Less Likely ↔ More Likely



Multi-Domain Dialogue State Tracking

- A full representation of the system's belief of the user's goal at any point during the dialogue
- Used for making API calls

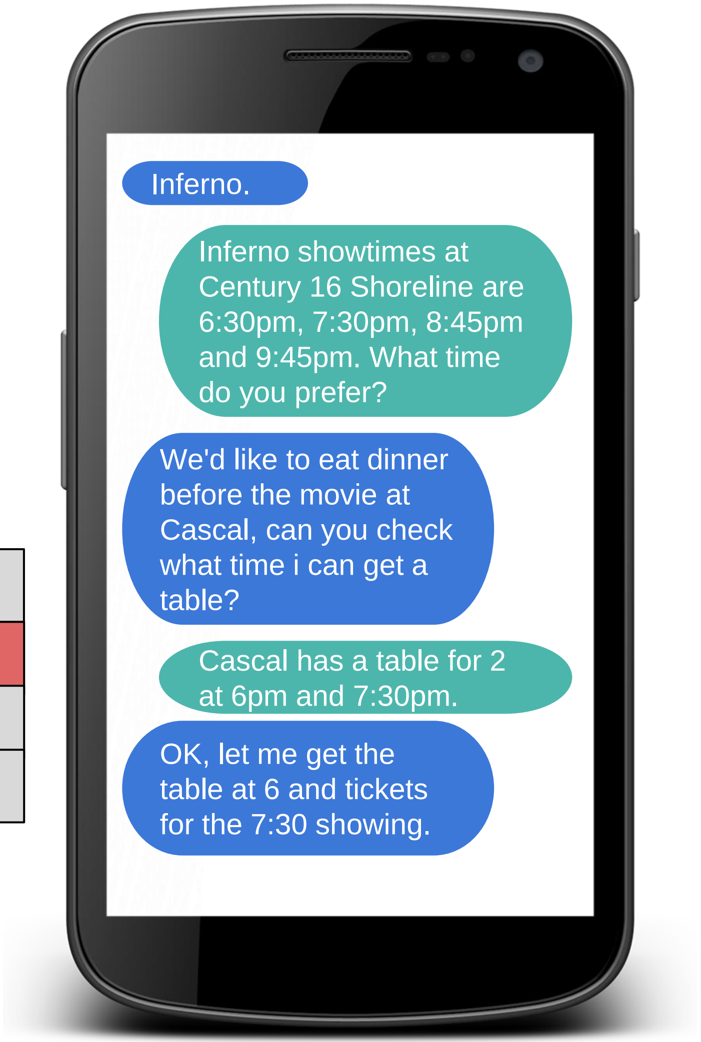
Movies

Date	11/15/17			
Time	6:30 pm	7:30 pm	8:45 pm	9:45 pm
#People	2			
Theater	Century 16 Shoreline			
Movie	Inferno			

Restaurants

Date	11/15/17		
Time	6:00 pm	6:30 pm	7:00 pm
Restaurant	Cascal		
#People	2		

Less Likely — More Likely



Discriminative DST – Single Turn

Data

- Observations labeled w/ dialogue state

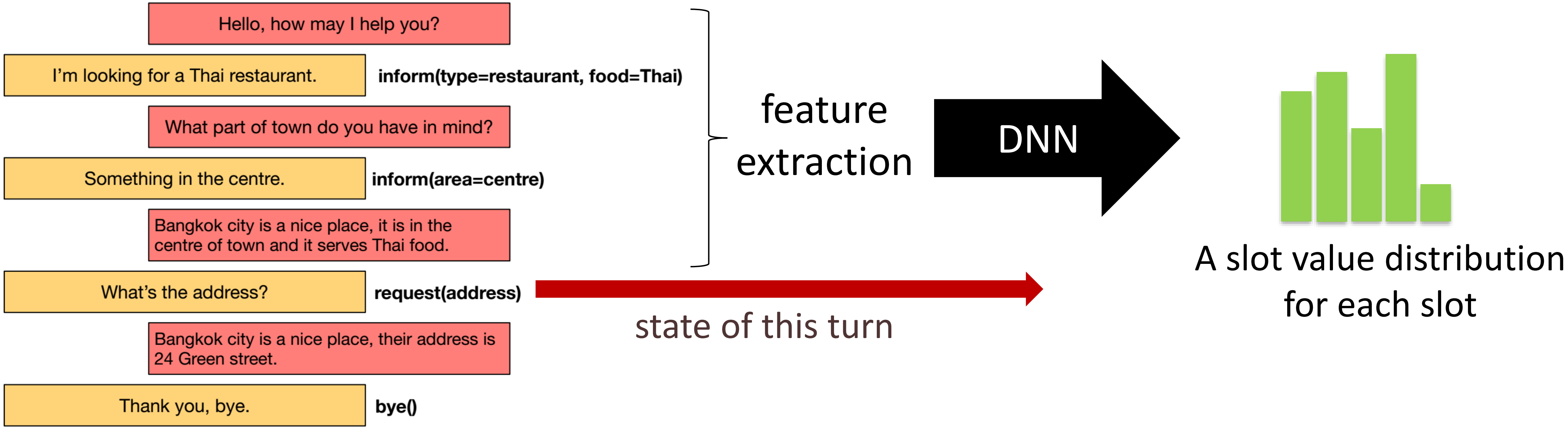
Model

- Neural networks
- Ranking models

Prediction

- Distribution over dialogue states
  – Dialogue State Tracking

DNN for DST



multi-turn conversation

Discriminative DST – Multiple Turns

Data

- Sequence of observations labeled w/ dialogue states

Model

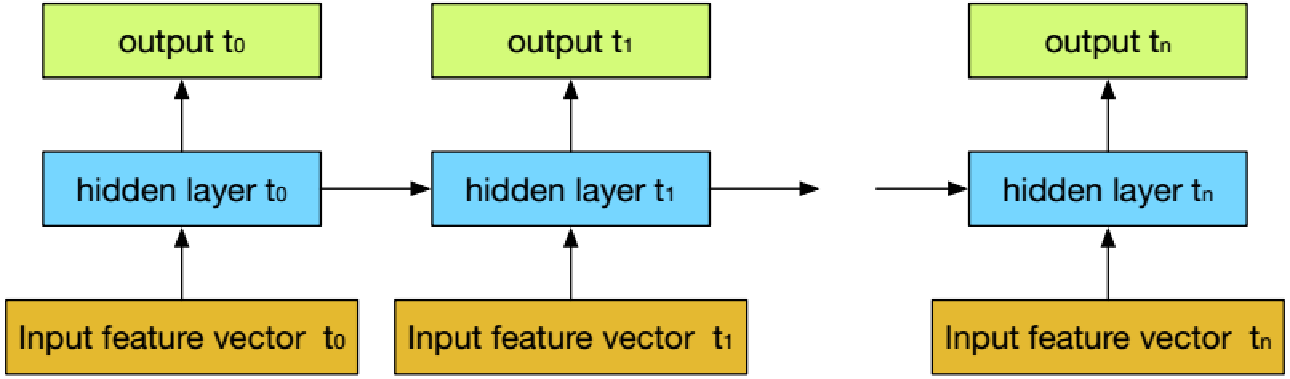
- Sequential models
  - Recurrent neural networks (RNN)

Prediction

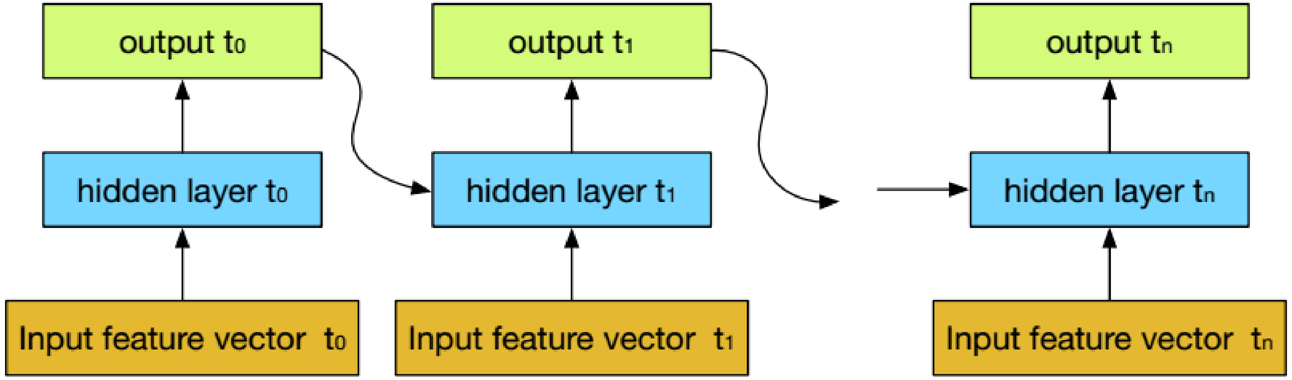
- Distribution over dialogue states
  – Dialogue State Tracking

Recurrent Neural Network (RNN)

- Elman-type

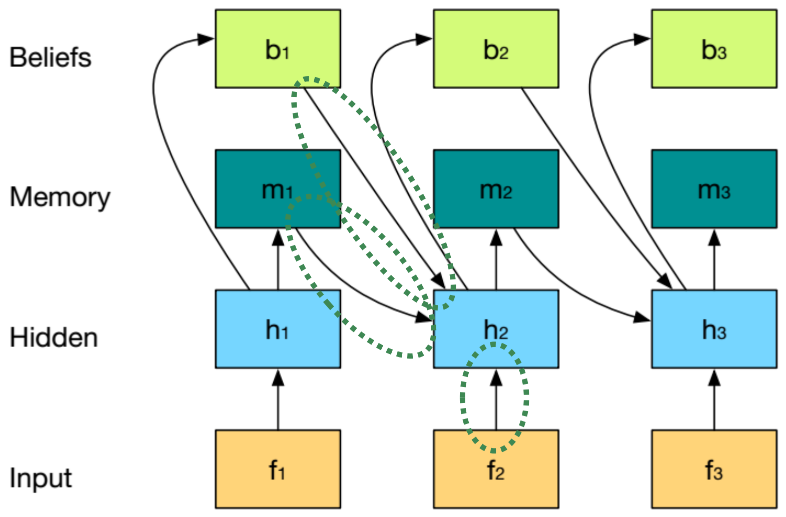


- Jordan-type

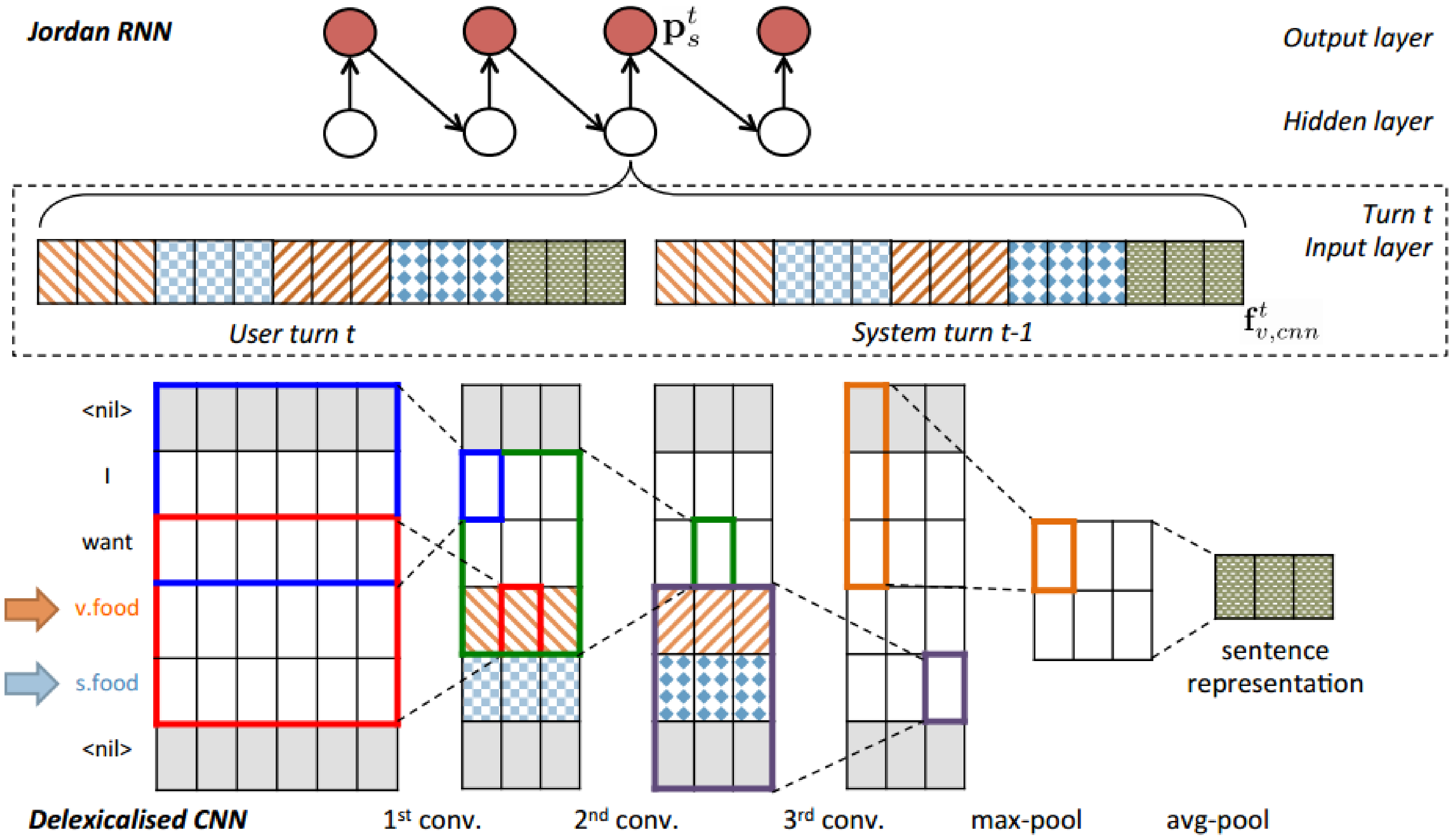


RNN-Based DST

- Idea: internal memory for representing dialogue context
  - Input
    - most recent dialogue turn
    - last machine dialogue act
    - dialogue state
    - memory layer
  - Output
    - update its internal memory
    - distribution over slot values



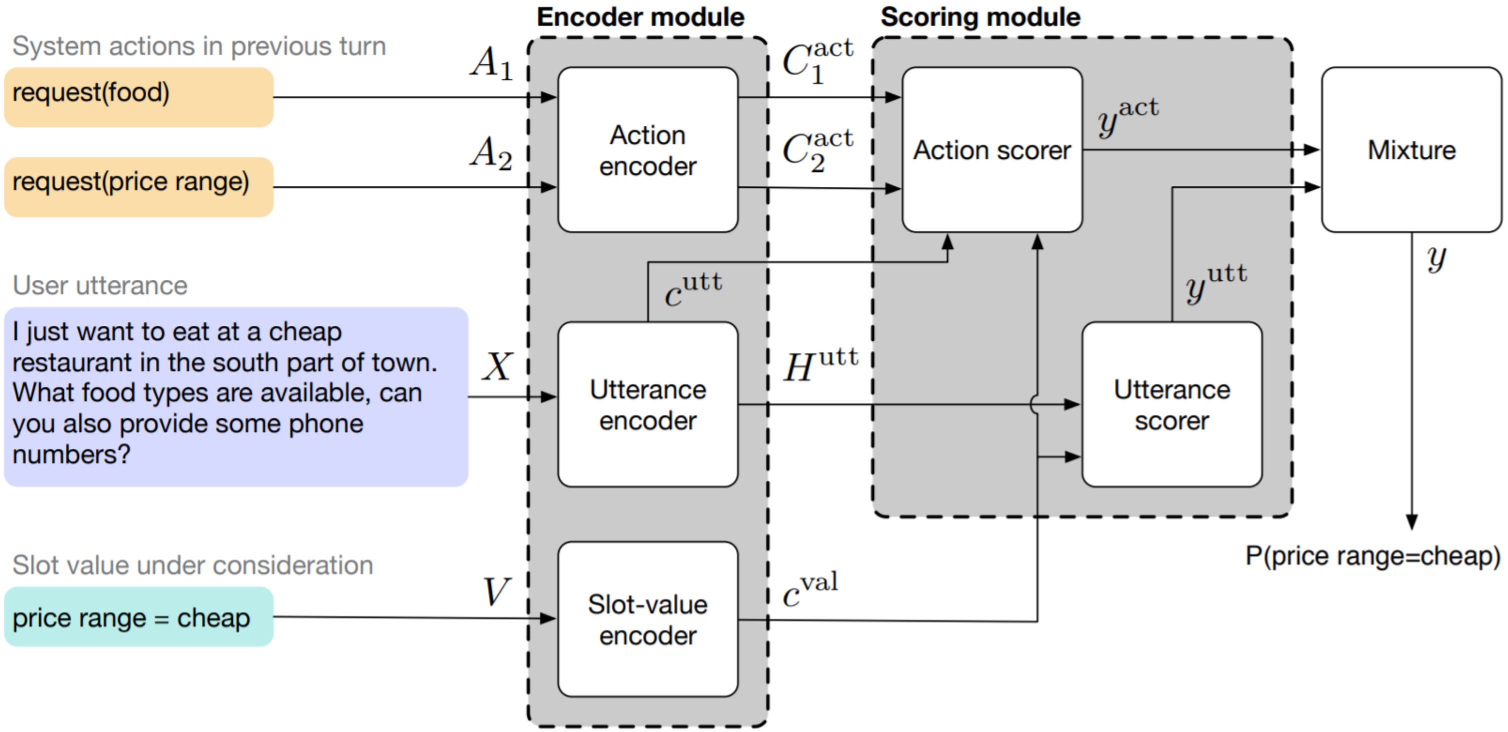
RNN-CNN DST (Mrkšić+, 2015)



(Figure from Wen et al, 2016)

Global-Locally Self-Attentive DST (Zhong+, 2018)

- More advanced encoder
  - Global modules share parameters for all slots
  - Local modules learn slot-specific feature representations



Generative DST

- Generating the state as a sequence (Lei+, 2018) or dialogue state updates (Lin+, 2020)

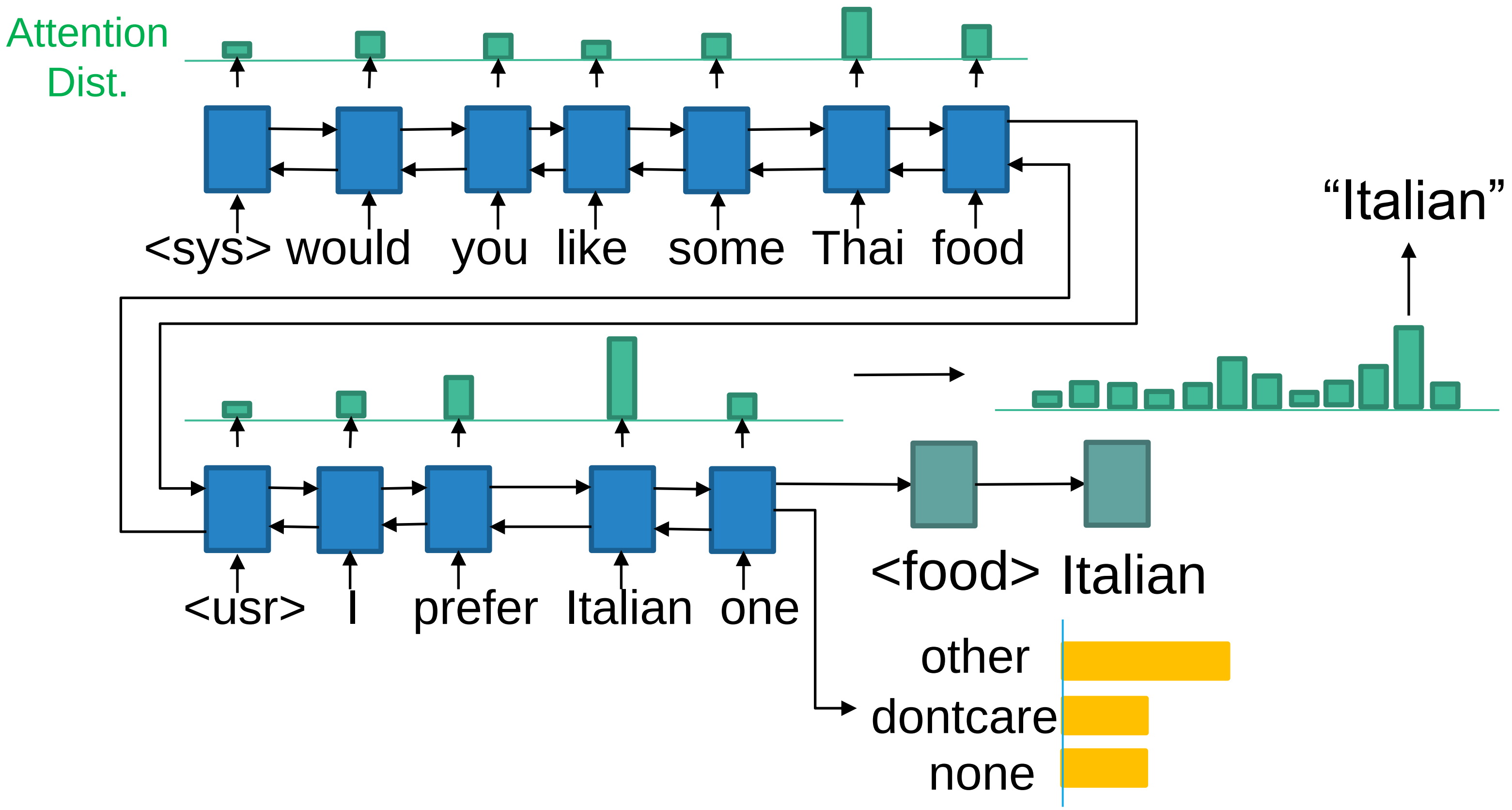
  (Dialogue history) ⇒ (slot1=val,slot2=val …)

- Given a dialogue and a slot, generate the value of the slot (Wu+, 2019; Gao+, 2019; Ren+, 2019; Zhou & Small, 2019; Kim+, 2019; Le+, 2020) ⇒ requires multiple forwards

  (Dialogue history, slot1) ⇒ val

Handling Unknown Slot Values (Xu & Hu, 2018)

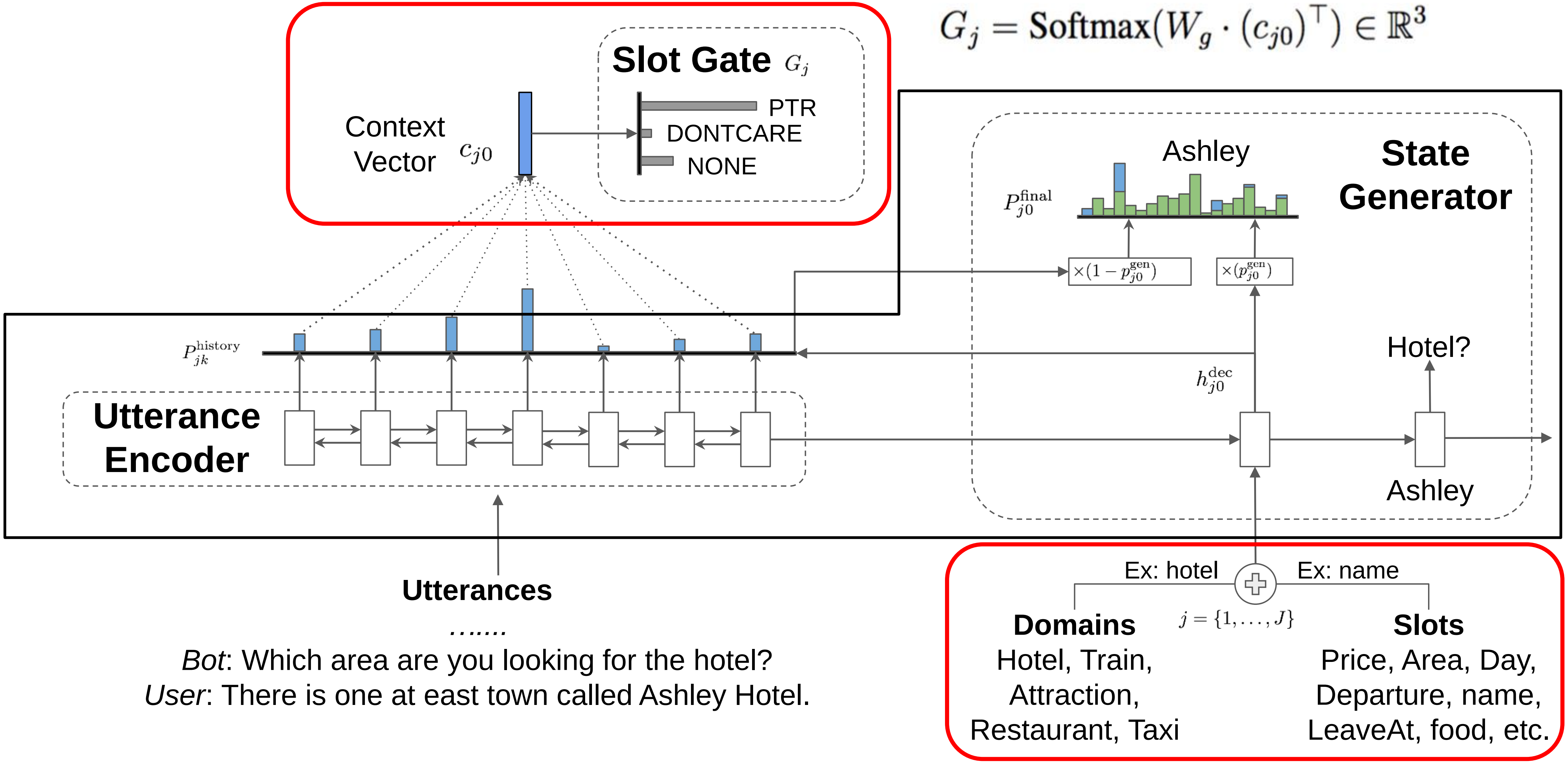
- Issue: fixed value sets in DST



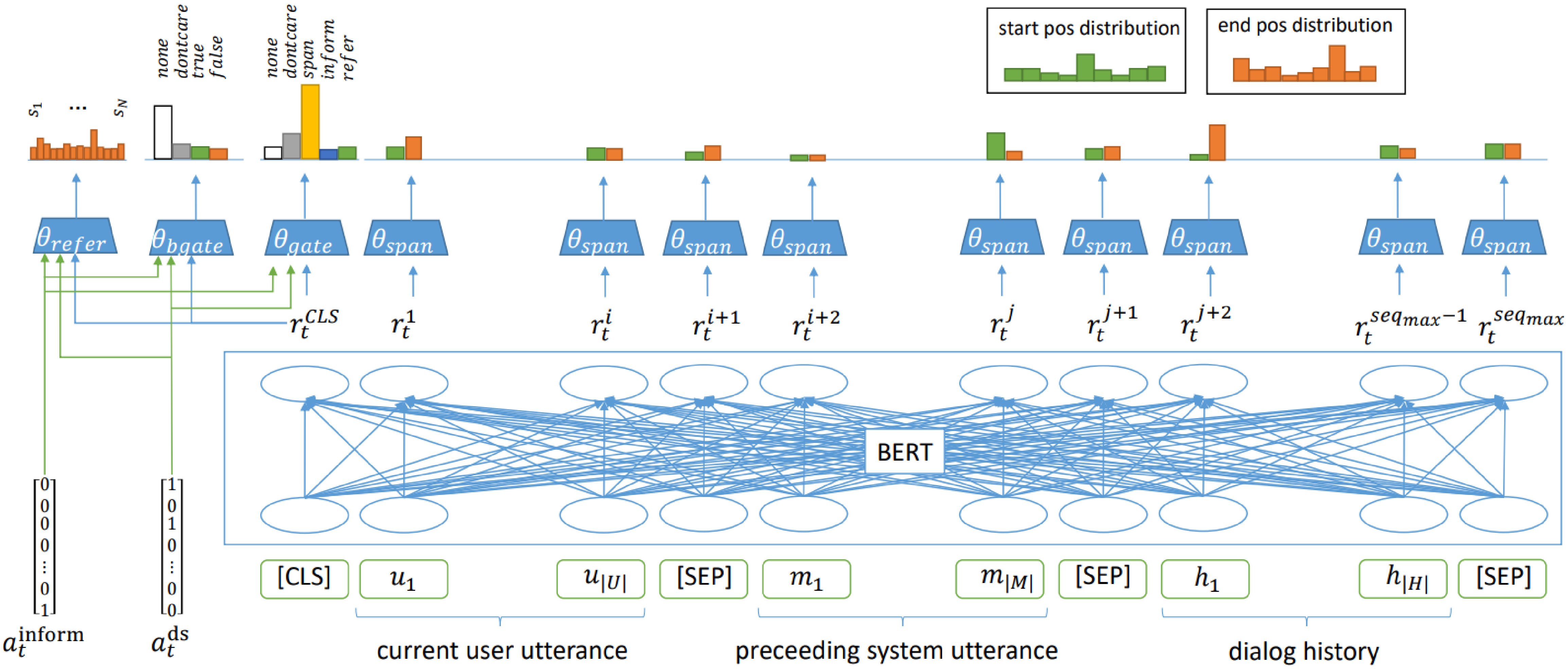
Pointer networks for generating unknown values

TRADE: Transferable DST (Wu+, 2019)

$$G_j = \mathrm{Softmax}(W_g \cdot (c_{j0})^\top) \in \mathbb{R}^3$$



TripPy: Handling OOV & Rare Values (Heck+, 2020)



DST Evaluation

Input Dialogue:
USER: Can you help me book a 5-star hotel on Sunday?
SYSTEM: For how many people?
USER: For two people, thanks!

Output Dialogue State:
Hotel_Book (star=5, day=sunday)

Hotel_Book (star=5, day=sunday, people_num=2)

⇒ Slot Acc / Joint Acc

- Dialogue State Tracking Challenges
  - DSTC2-3, human-machine
  - DSTC4-5, human-human
  - DSTC8, human-machine
- Metric
  - Tracked state accuracy with respect to user goal
  - Recall/Precision/F-measure individual slots

Dialog State Tracking Challenge (DSTC)

(Williams et al. 2013, Henderson et al. 2014, Henderson et al. 2014, Kim et al. 2016, Kim et al. 2016)

Challenge	Type	Domain	Data Provider	Main Theme
DSTC1	Human-Machine	Bus Route	CMU	Evaluation Metrics
DSTC2	Human-Machine	Restaurant	U. Cambridge	User Goal Changes
DSTC3	Human-Machine	Tourist Information	U. Cambridge	Domain Adaptation
DSTC4	Human-Human	Tourist Information	I2R	Human Conversation
DSTC5	Human-Human	Tourist Information	I2R	Language Adaptation

DSTC4-5

- Type: Human-Human
- Domain: Tourist Information

{Topic: Accommodation; NAME: InnCrowd Backpackers Hostel; GuideAct: REC; TouristAct: ACK}

Guide: Let's try this one, okay?

Tourist: Okay.

Guide: It's InnCrowd Backpackers Hostel in Singapore. If you take a dorm bed per person only twenty dollars. If you take a room, it's two single beds at fifty nine dollars.

Tourist: Um. Wow, that's good.

Guide: Yah, the prices are based on per person per bed or dorm. But this one is room. So it should be fifty nine for the two room. So you're actually paying about ten dollars more per person only.

Tourist: Oh okay. That's- the price is reasonable actually. It's good.

{Topic: Accommodation; Type: Hostel; Pricerange: Cheap; GuideAct: ACK; TouristAct: REQ}

Tourist: Can you give me some uh- tell me some cheap rate hotels, because I'm planning just to leave my bags there and go somewhere take some pictures.

Guide: Okay. I'm going to recommend firstly you want to have a backpack type of hotel, right?

Tourist: Yes. I'm just gonna bring my backpack and my buddy with me. So I'm kinda looking for a hotel that is not that expensive. Just gonna leave our things there and, you know, stay out the whole day.

Guide: Okay. Let me get you hm hm. So you don't mind if it's a bit uh not so roomy like hotel because you just back to sleep.

Tourist: Yes. Yes. As we just gonna put our things there and then go out to take some pictures.

Guide: Okay, um-

Tourist: Hm.

Multi-Domain DST Data

- MultiWoZ 2.0 ⇒ 2.1 ⇒ 2.2 ⇒ 2.3 ⇒ ……

act type	inform* / request* / select[123] / recommend/[123] / not found[123] request booking info[123] / offer booking[1235] / inform booked[1235] / decline booking[1235] welcome* /greet* / bye* / reqmore*
slots	address* / postcode* / phone* / name[1234] / no of choices[1235] / area[123] / pricerange[123] / type[123] / internet[2] / parking[2] / stars[2] / open hours[3] / departure[45] destination[45] / leave after[45] / arrive by[45] / no of people[1235] / reference no.[1235] / trainID[5] / ticket price[5] / travel time[5] / department[7] / day[1235] / no of days[123]

- SGD: natural language described schema for better scalability

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

Intents
name: *"MakePayment"*
description: *"Send money to your contact"*
required_slots: [*"amount"*, *"contact_name"*]
optional_slots: [*"account_type"* = *"in-app balance"*]

name: *"RequestPayment"*
description: *"Request money from a contact"*
required_slots: [*"amount"*, *"contact_name"*]

Slots
name: *"account_type"* categorical: True
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"*

MultiWOZ 2.1 Leaderboard

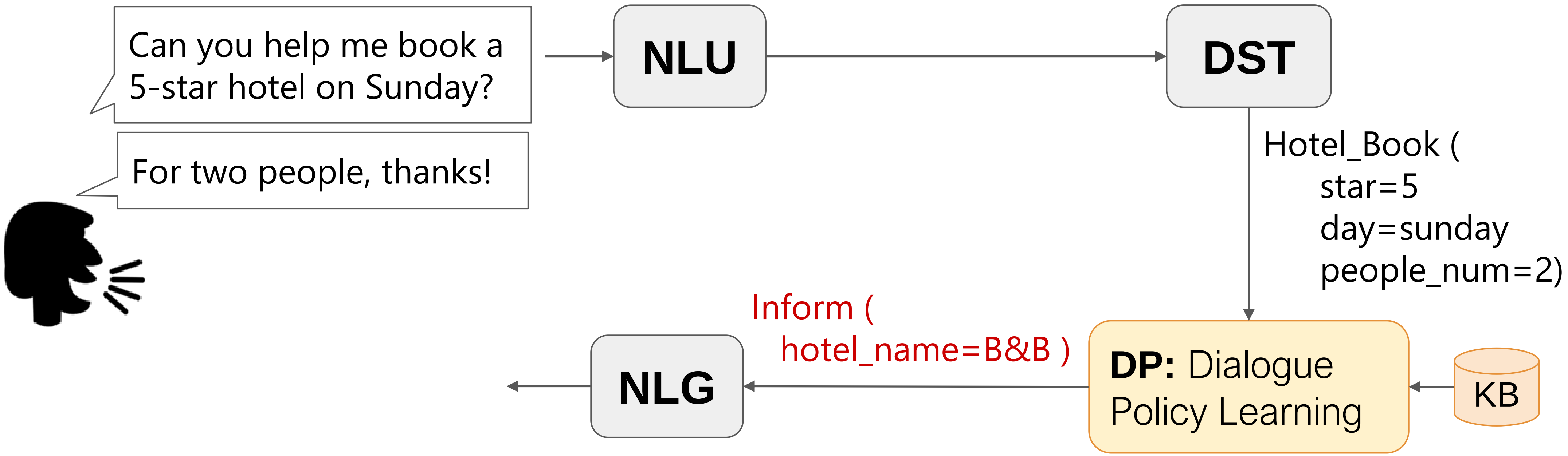
Rank	Model	Joint Acc	Paper	Code	Result	Year
1	CHAN-DST	58.55	A Contextual Hierarchical Attention Network with Adaptive Objective for Dialogue State Tracking			2020
2	SimpleTOD	55.76	A Simple Language Model for Task-Oriented Dialogue			2020
3	Transformer-DST	55.35	Jointly Optimizing State Operation Prediction and Value Generation for Dialogue State Tracking			2020
4	TripPy	55.30	TripPy: A Triple Copy Strategy for Value Independent Neural Dialog State Tracking			2020
5	SST	55.23	Schema-Guided Multi-Domain Dialogue State Tracking with Graph Attention Neural Networks			2020
6	Graph-DST	53.85	Multi-Domain Dialogue State Tracking based on State Graph			2020
7	DS-Picklist	53.30	Find or Classify? Dual Strategy for Slot-Value Predictions on Multi-Domain Dialog State Tracking			2019
8	CSFN-DST + BERT	52.88	Efficient Context and Schema Fusion Networks for Multi-Domain Dialogue State Tracking			2020
9	SOM-DST	52.57	Efficient Dialogue State Tracking by Selectively Overwriting Memory			2019
10	DSTQA	51.17	Multi-domain Dialogue State Tracking as Dynamic Knowledge Graph Enhanced Question Answering			2019
11	NADST	49.04	Non-Autoregressive Dialog State Tracking			2020
12	TRADE	45.60	MultiWOZ 2.1: A Consolidated Multi-Domain Dialogue Dataset with State Corrections and State Tracking Baselines			2019

61

Dialogue Policy Learning

Modular Task-Oriented Dialogue Systems

Dialogue Policy Learning



- DP decides the system action for interacting with users based on dialogue states.
  - Input: dialogue state + KB results
  - Output: system action (speech-act + slot-value pairs)

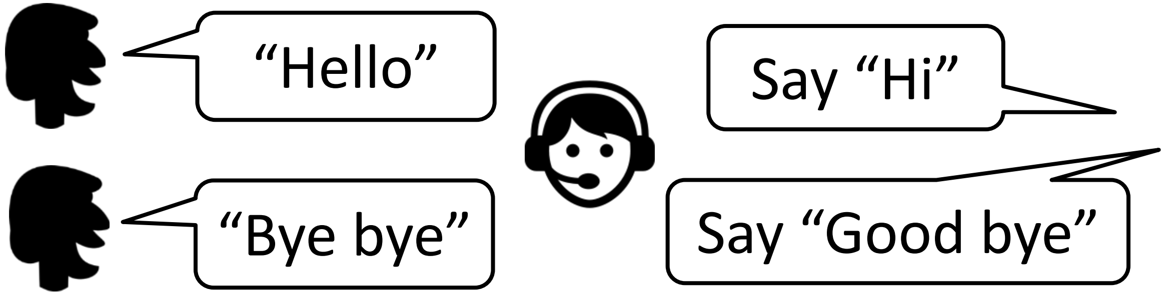
Dialogue Policy Learning

Hello, how may I help you?	greeting ()
I’m looking for a Thai restaurant.	request (restaurant; foodtype=Thai)
What part of town do you have in mind?	request (area)
Something in the centre.	inform (area=centre)
Bangkok city is a nice place, it is in the centre of town and it serves Thai food.	inform (restaurant=Bangkok city, area=centre of town, foodtype=Thai)
What’s the address?	request (address)
Bangkok city is a nice place, their address is 24 Green street.	inform (address=24 Green street)
Thank you, bye.	bye ()

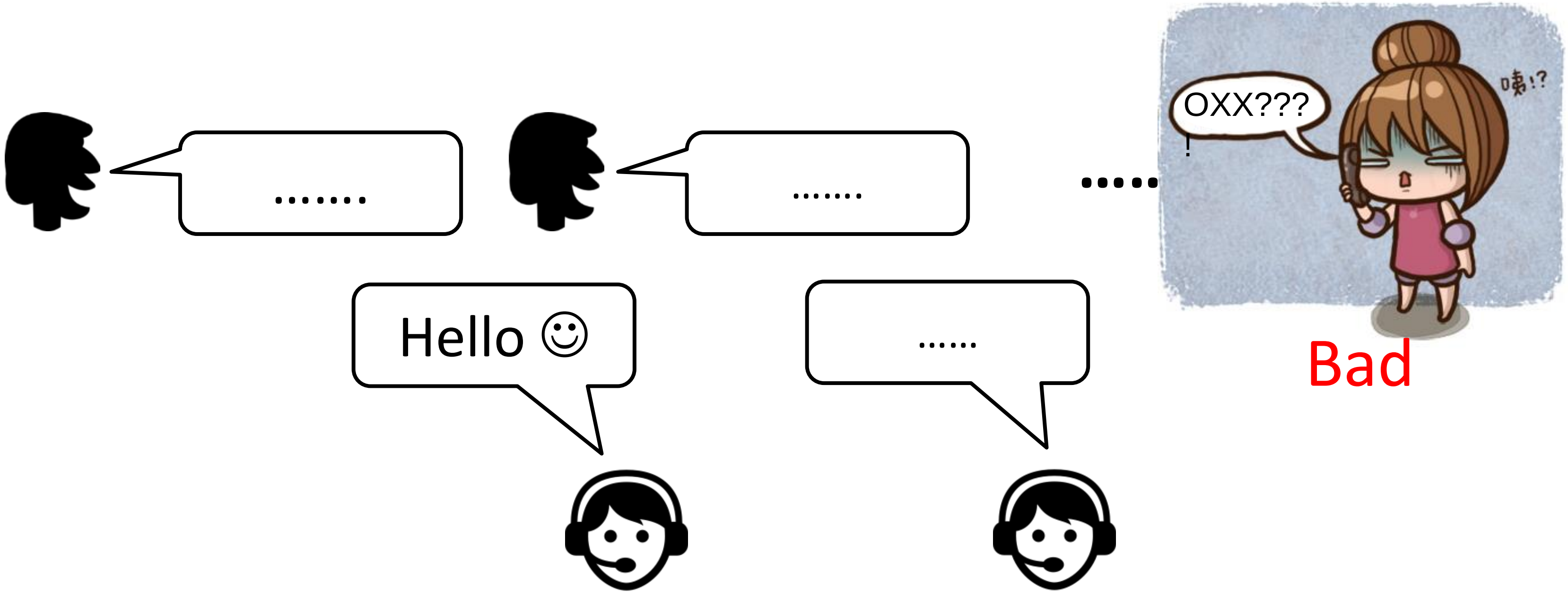
Supervised v.s. Reinforcement

- Supervised

Learning from teacher



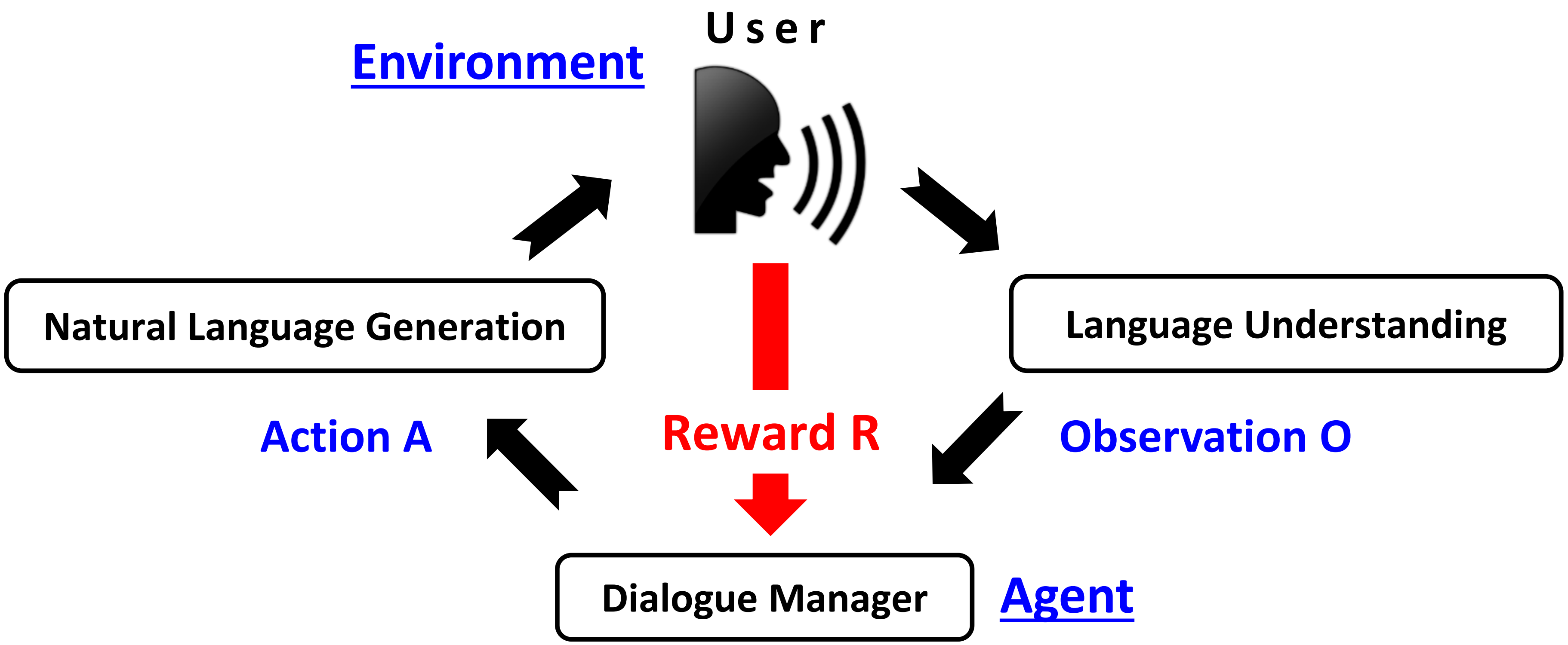
- Reinforcement



Learning from critics

Dialogue Policy Optimization

- Dialogue management in a RL framework



Select the best action that **maximizes the future reward**

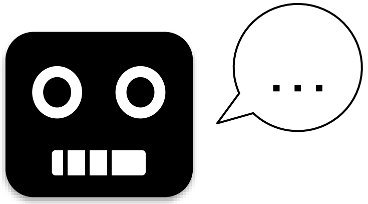
Reward for RL ≅ Evaluation for System

- Dialogue is a special RL task
  - Human involves in interaction and rating (evaluation) of a dialogue
  - Fully human-in-the-loop framework
- Rating: correctness, appropriateness, and adequacy

- Expert rating	high quality, high cost
- User rating	unreliable quality, medium cost
- Objective rating	Check desired aspects, low cost

Dialogue Reinforcement Learning Signal

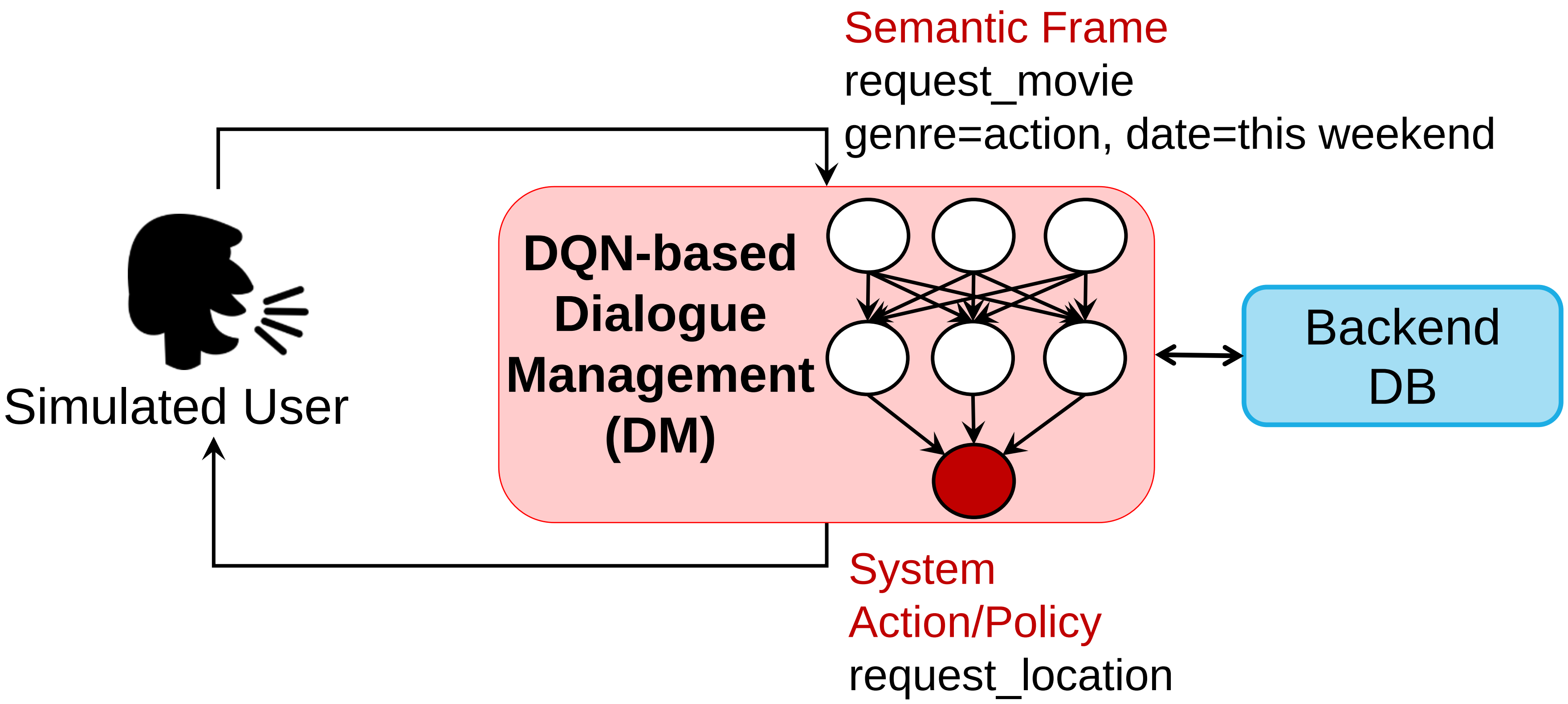
- Typical reward function
  - -1 for per turn penalty
  - Large reward at completion if successful
- Typically requires domain knowledge
  - ✔ Simulated user
  - ✔ Paid users (Amazon Mechanical Turk)
  - ✘ Real users



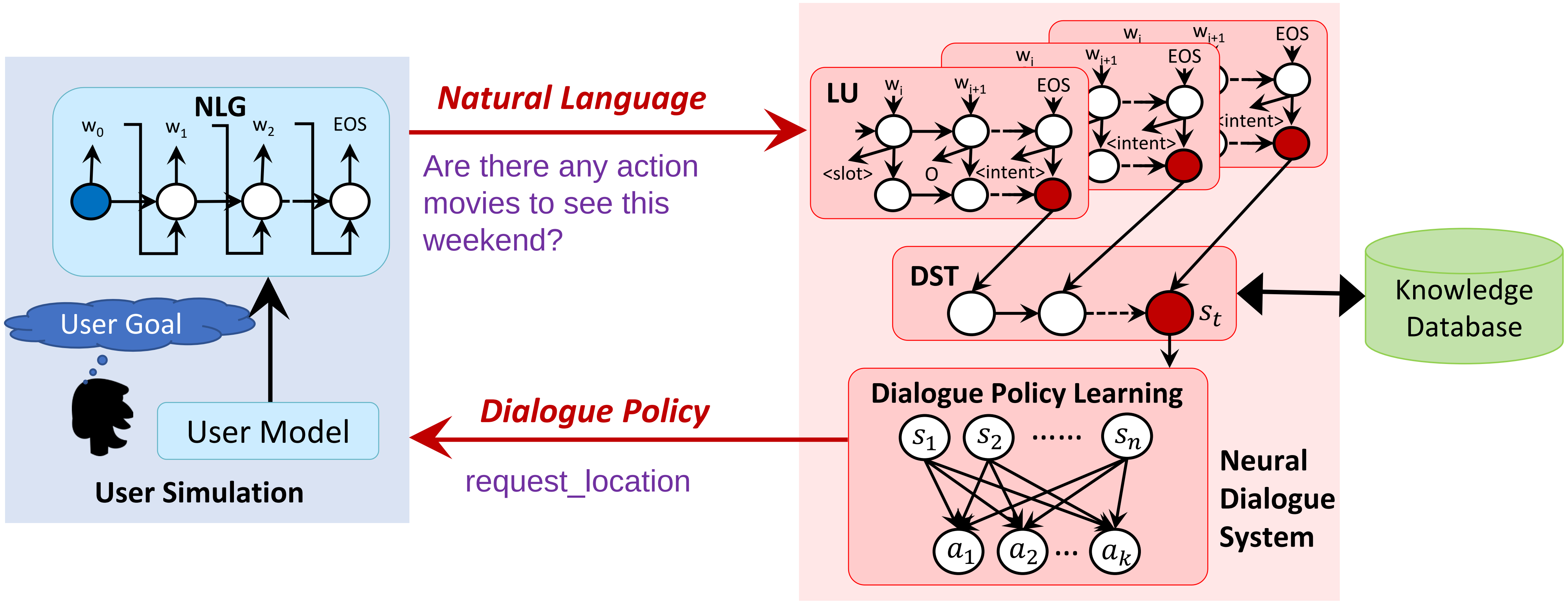
The user simulator is usually required for dialogue system training before deployment

Neural Dialogue Manager (Li et al., 2017)

- Deep Q-network for training dialogue policy
  - Input: current semantic frame observation, database returned results
  - Output: system action



E2E Task-Completion Bot (TC-Bot) (Li et al., 2017)



Idea: SL for each component and RL for end-to-end training the neural dialogue system

Reinforcement Learning Agent (Li et al., 2017)

User goal – ***Two tickets for "Deadpool" tomorrow 9PM at AMC Pacific Place 11 theater, Seattle.***

RULE BASED SYSTEM
User: Do you have any ticket for movie 'Deadpool' at 9:00 pm?
Agent: Which theater do you want?
User: AMC pacific place 11 theater, please!
Agent: How many tickets do you want?
User: Two!
Agent: What date?
User: Tomorrow!
Agent: Which city do you want?
User: Seattle!
Agent: Do you have any specification for zip code?
User: No, I do not care.
Agent: Do you have any distance constraint?
User: No, I do not care.
Agent: How about video format? 2d? Imax 3d?
User: Any kind is okay.
Agent: Do you have any constraint to price?
User: No!
Agent: Okay, I will issue 2 tickets for you, tomorrow 9:00 pm at AMC pacific place 11 theater, Seattle, movie 'Deadpool'.
User: Okay, t
Agent: Thanks

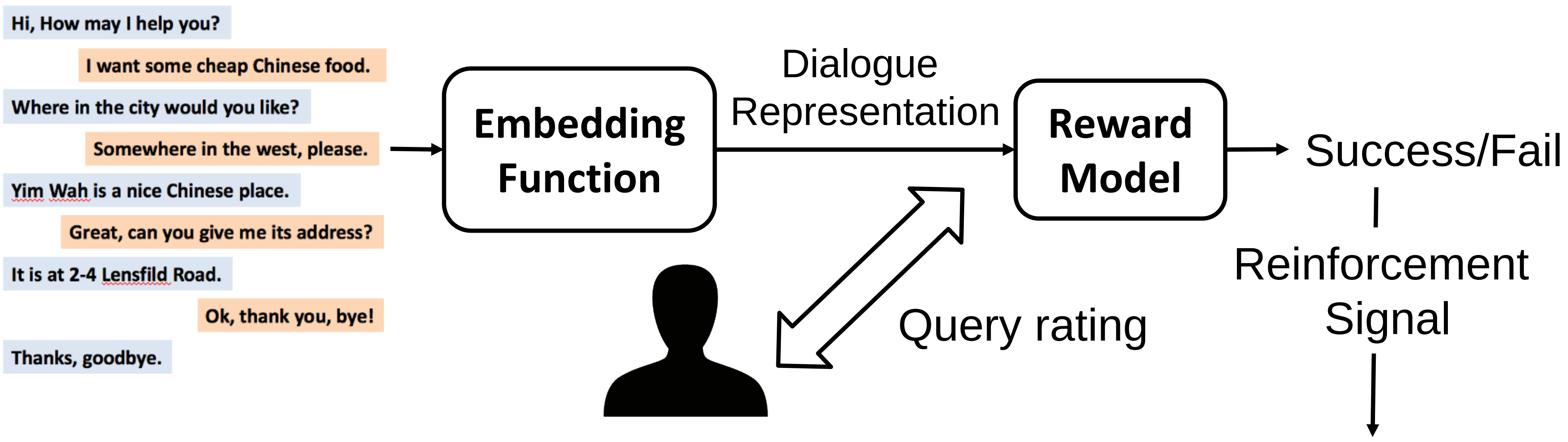
Issue: no notion about what requests can be skipped

REINFORCEMENT LEARNING SYSTEM
User: I want two tickets for movie 'Deadpool'.
Agent: What is the starttime you prefer?
User: 9:00 PM
Agent: Which day?
User: Tomorrow
Agent: Which city?
User: Seattle
Agent: Which theater do you prefer?
User: AMC Pacific Place 11 theater.
Agent: Okay, I will issue 2 tickets for you, tomorrow 9:00 pm at AMC pacific place 11 theater, Seattle, movie 'Deadpool'.
User: Okay, thanks.
Agent: Thanks!

Skip the requests the user may not care about to improve efficiency

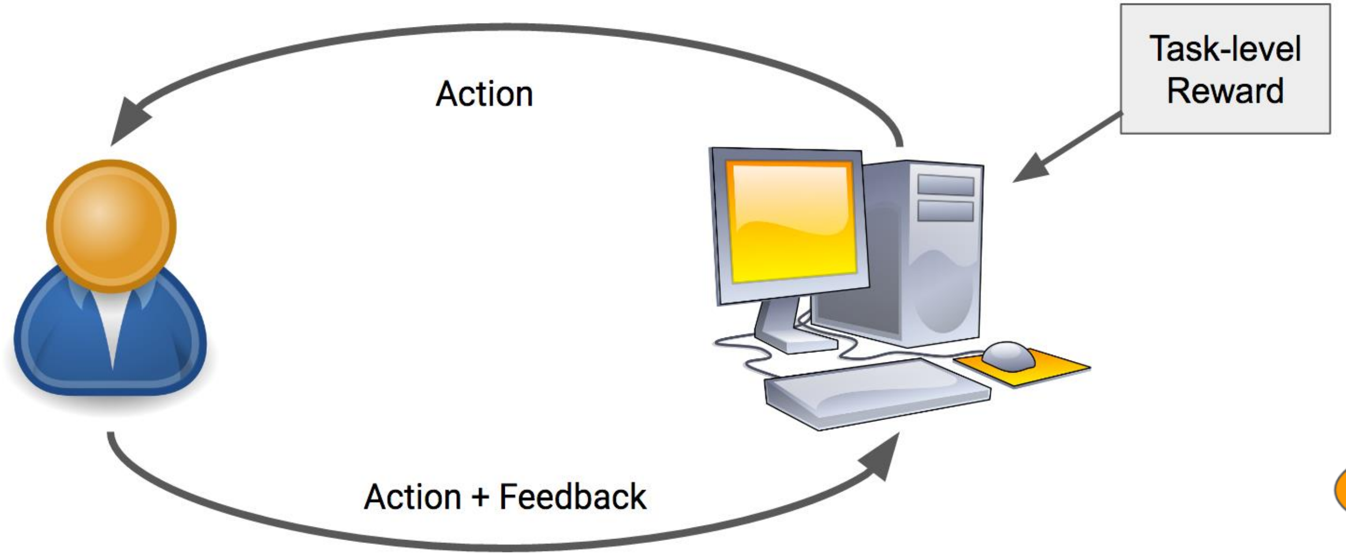
Online Training (Su+, 2015; Su+, 2016)

- Policy learning from real users
  - Infer reward directly from dialogues (Su et al., 2015)
  - User rating (Su et al., 2016)
- Reward modeling on user binary success rating



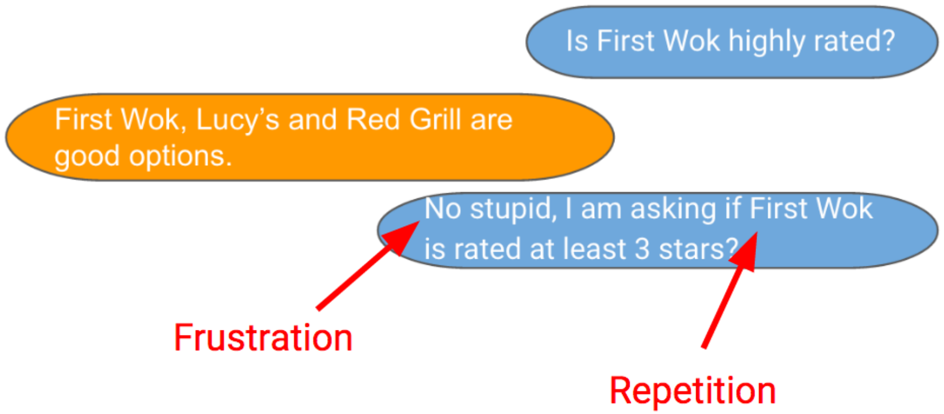
Interactive RL for DP (Shah+, 2016)

Immediate Feedback



Explicit

Implicit

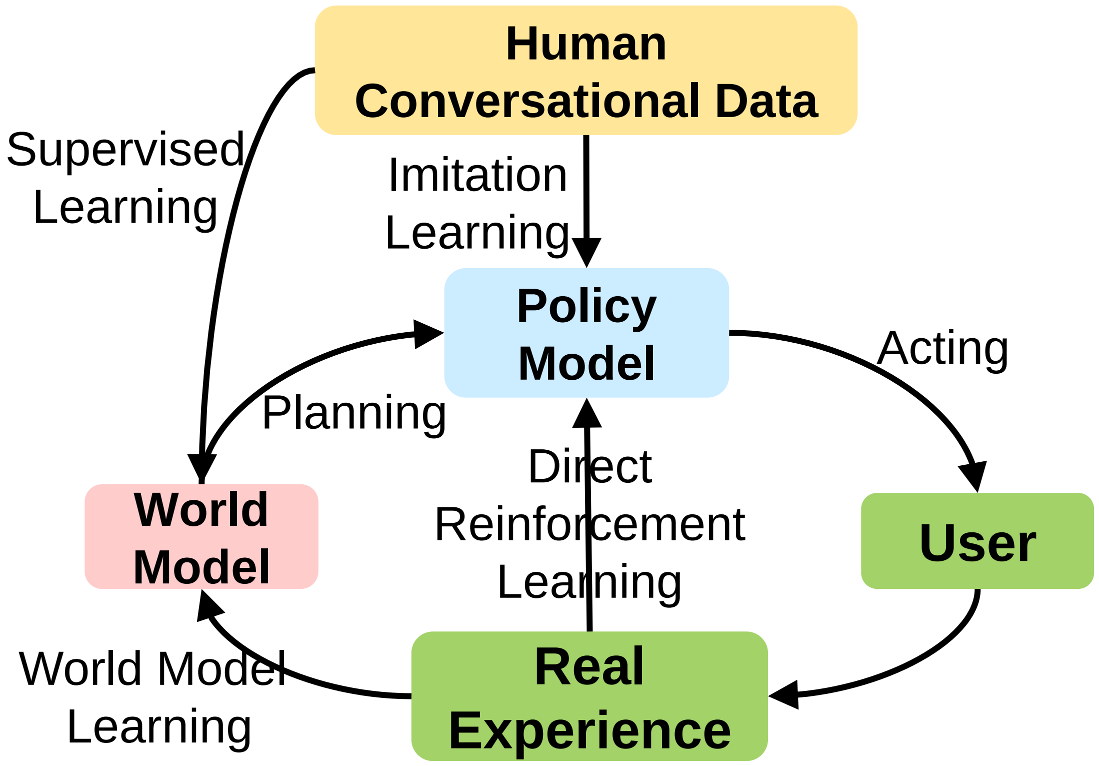


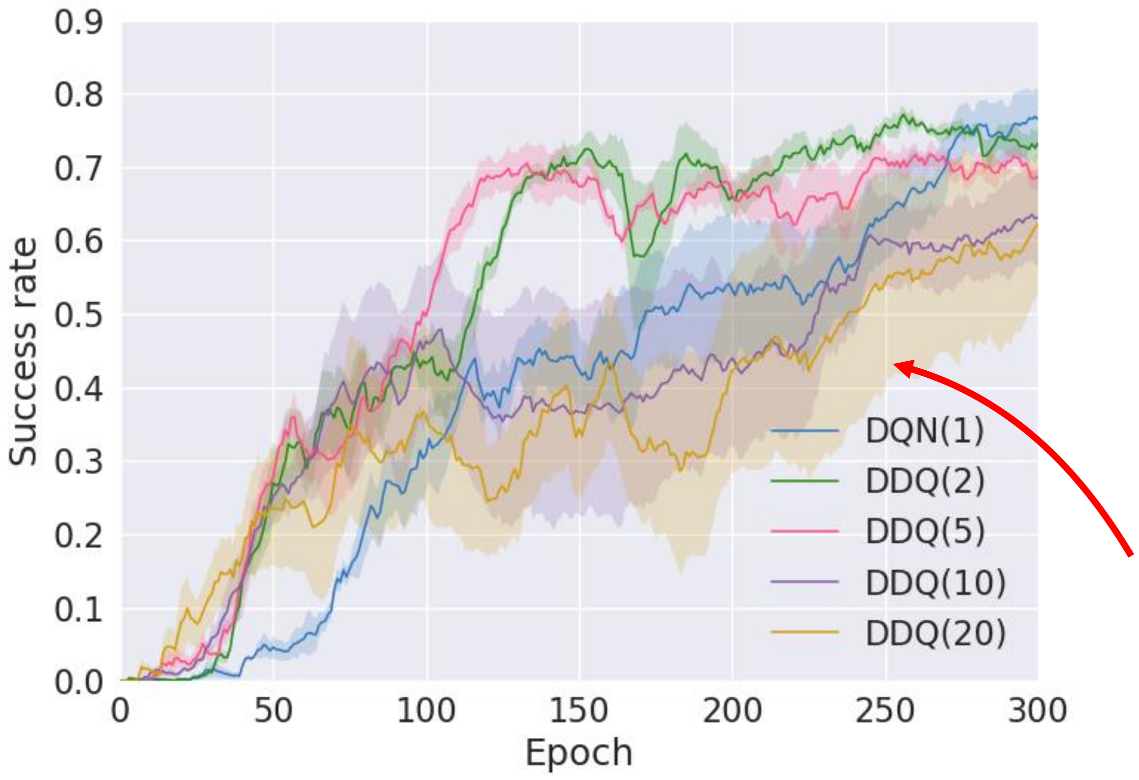
Use a third agent for providing interactive feedback to the policy

73

Planning – Deep Dyna-Q (Peng+, 2018)

- Idea: learning with real users with planning

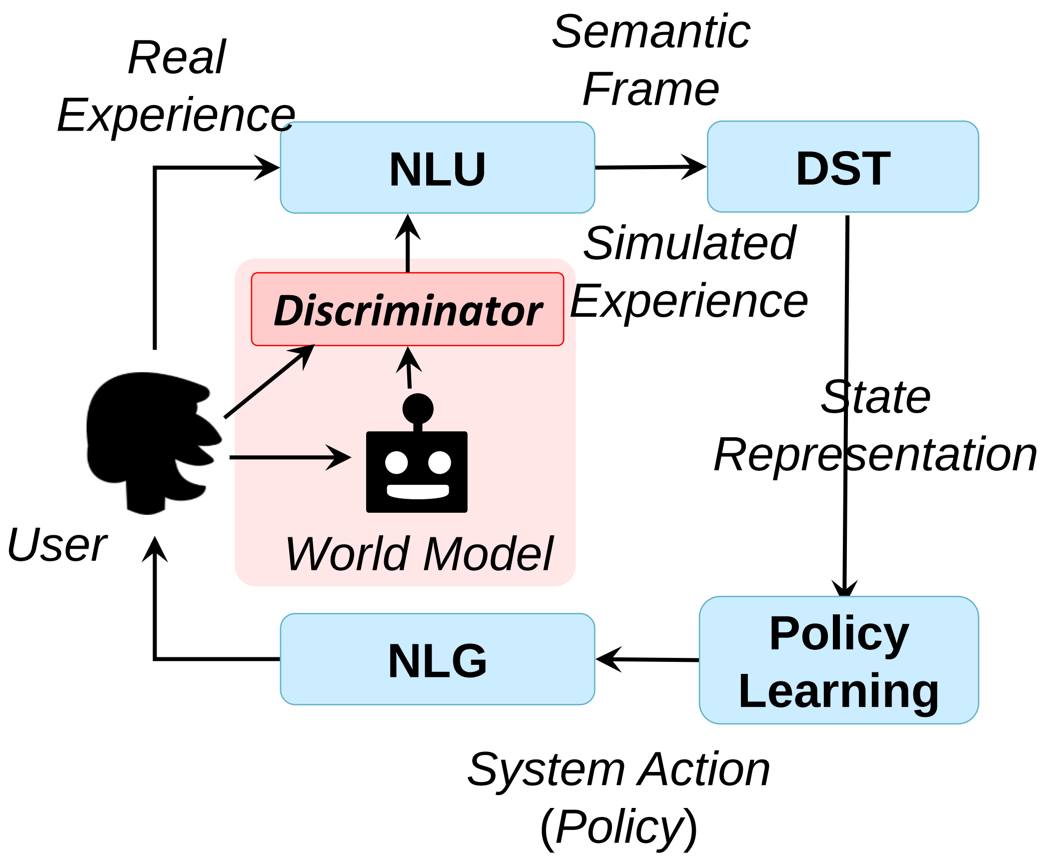


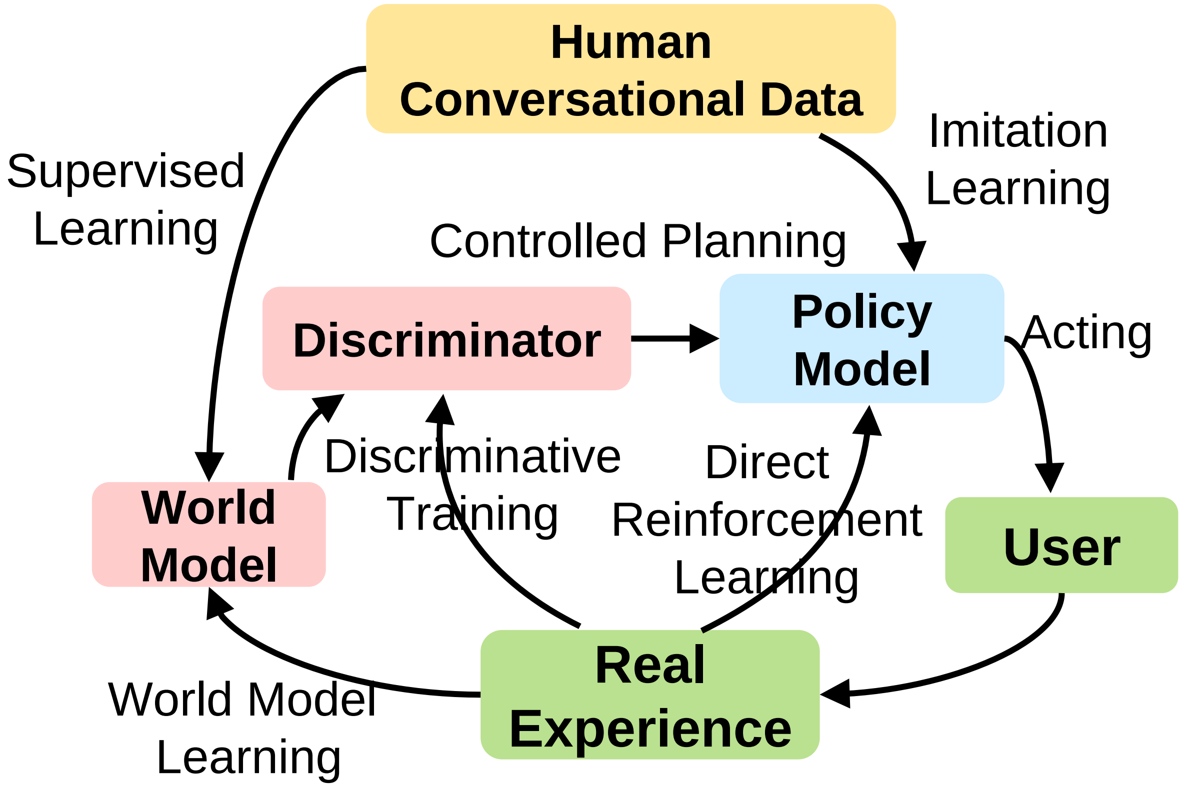


Policy learning suffers from the poor quality of fake experiences

Robust Planning – D3Q (Su+, 2018)

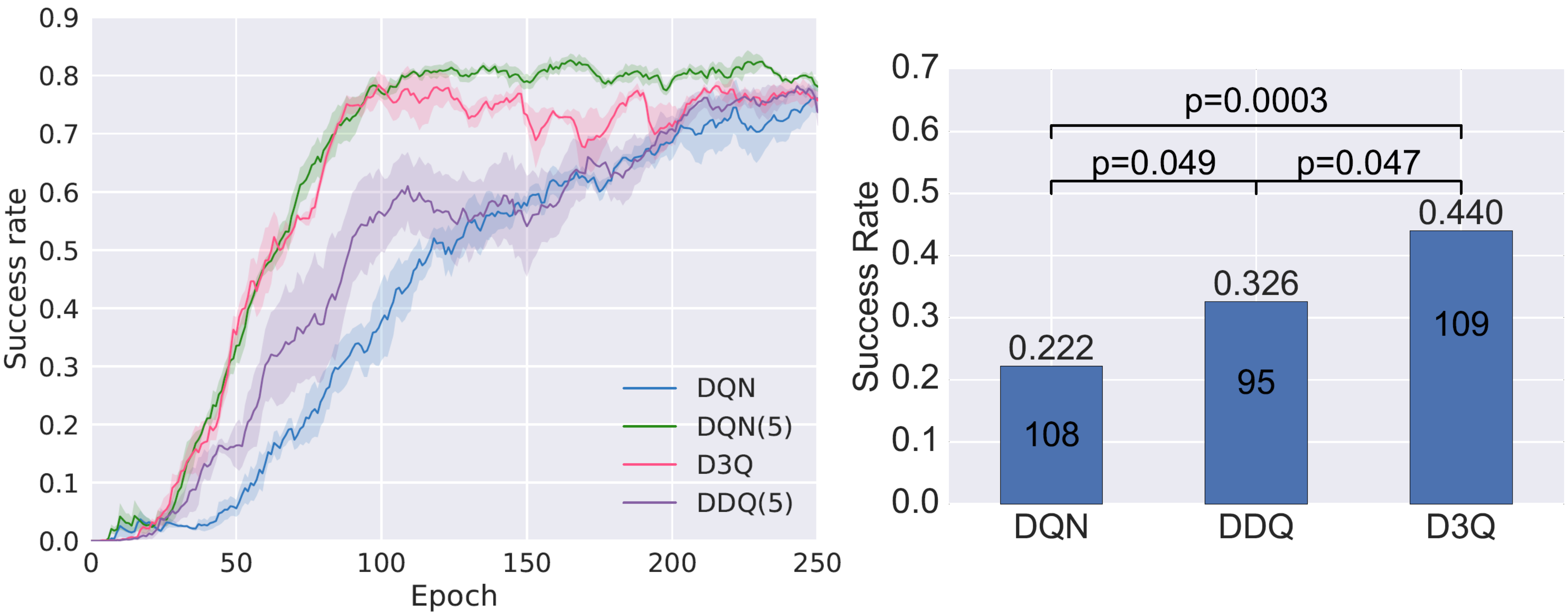
- Idea: add a *discriminator* to filter out the bad experiences





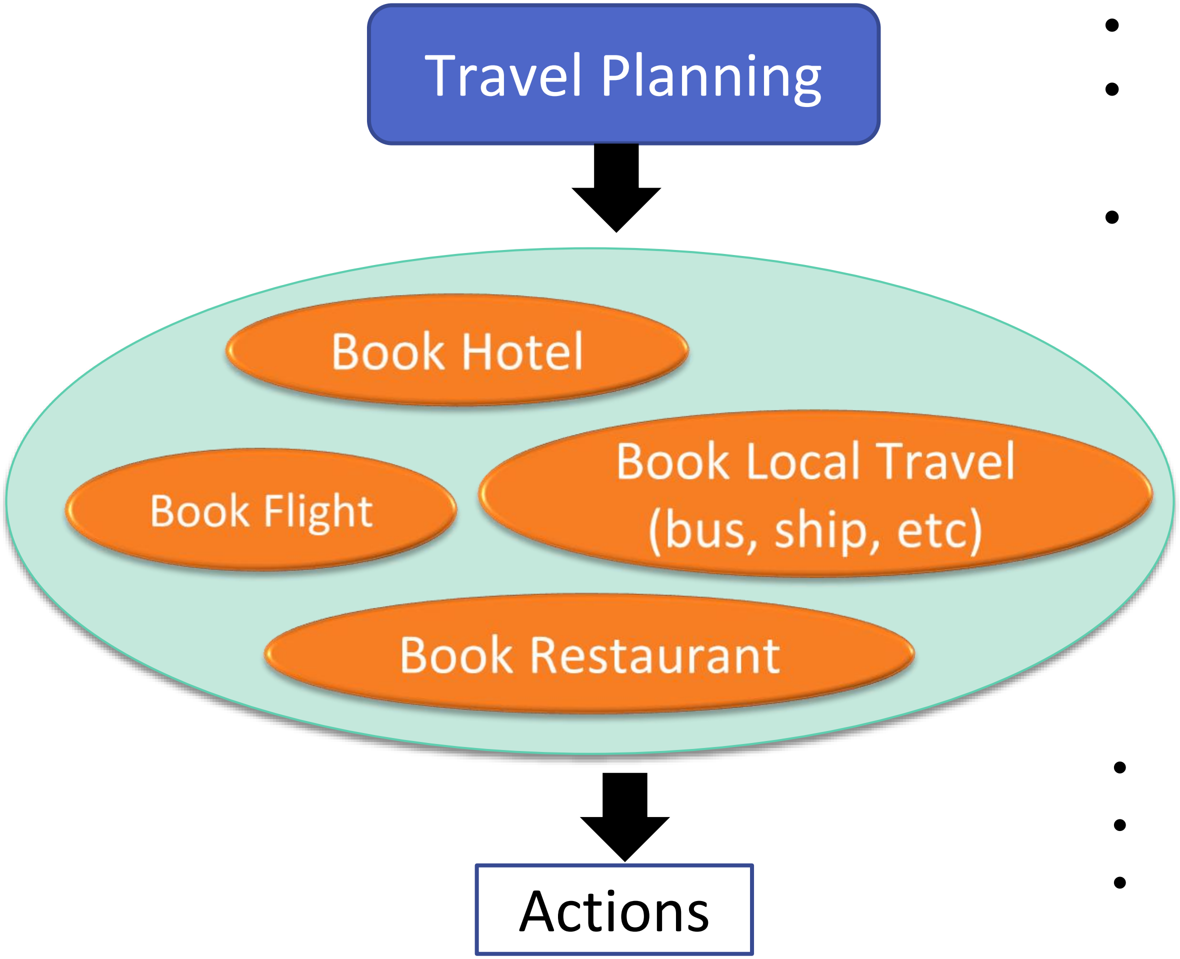
S.-Y. Su, X. Li, J. Gao, J. Liu, and Y.-N. Chen, "Discriminative Deep Dyna-Q: Robust Planning for Dialogue Policy Learning," (to appear) in *Proc. of EMNLP, 2018.*

Robust Planning – D3Q (Su+, 2018)



The policy learning is more robust and shows the improvement in human evaluation

Multi-Domain – Hierarchical RL (Peng+, 2017)



- Set of tasks that need to be fulfilled collectively!
- Build a DM for **cross-subtask constraints** (**slot constraints**)
- Temporally constructed goals

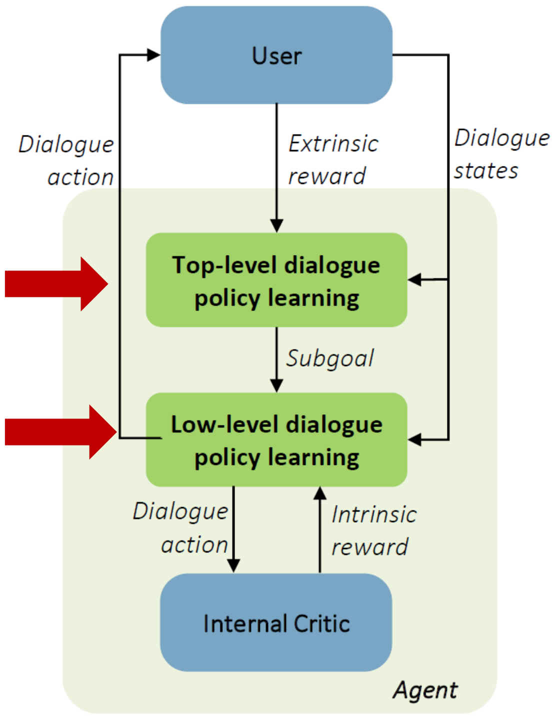
- hotel_check_in_time > departure_flight_time
- # flight_tickets = #people checking in the hotel
- hotel_check_out_time< return_flight_time

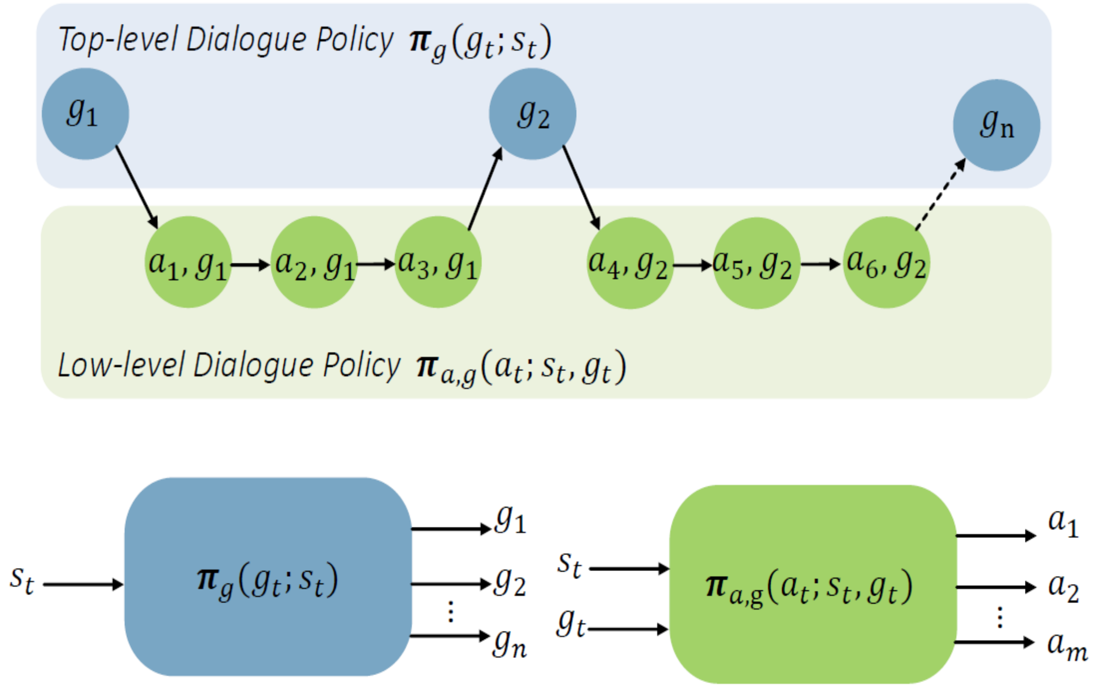
Multi-Domain – Hierarchical RL (Peng+, 2017)

- Model makes decisions over two levels: *meta-controller & controller*
- The *agent* learns these policies simultaneously
  - Policy of optimal sequence of goals to follow $\pi_g(g_t, s_t; \theta_1)$
  - Policy $\pi_{a,g}(a_t, g_t, s_t; \theta_2)$ for each sub-goal g_t

Meta-Controller

Controller





Multiple policies need to collaborate with each other for better multi-domain interactions

Dialogue Policy Evaluation

Dialogue State:
Hotel_Book (star=5, day=sunday, people_num=2)
KB State:
rest1=B&B

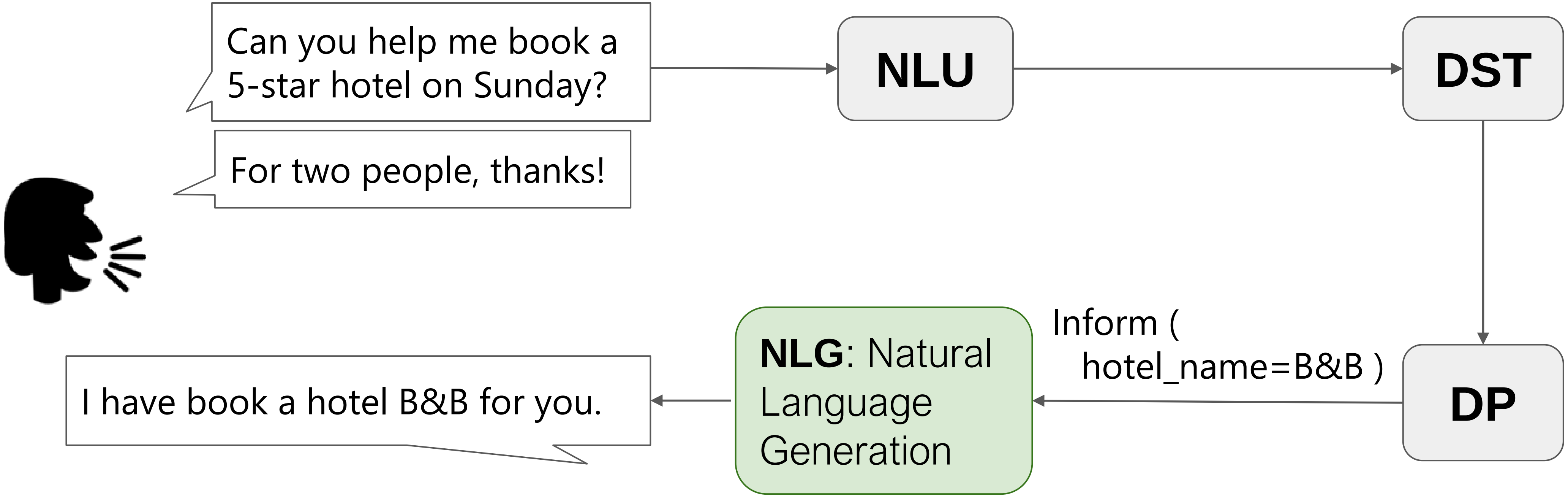
System Action:
inform (hotel_name=B&B)

- Metrics
  - Turn-level evaluation: system action accuracy
  - Dialogue-level evaluation: task success rate, reward

Natural Language Generation

Modular Task-Oriented Dialogue Systems

Natural Language Generation



- NLG is to map system actions to natural language responses.
  - Input: system speech-act + slot-value (optional)
  - Output: natural language response

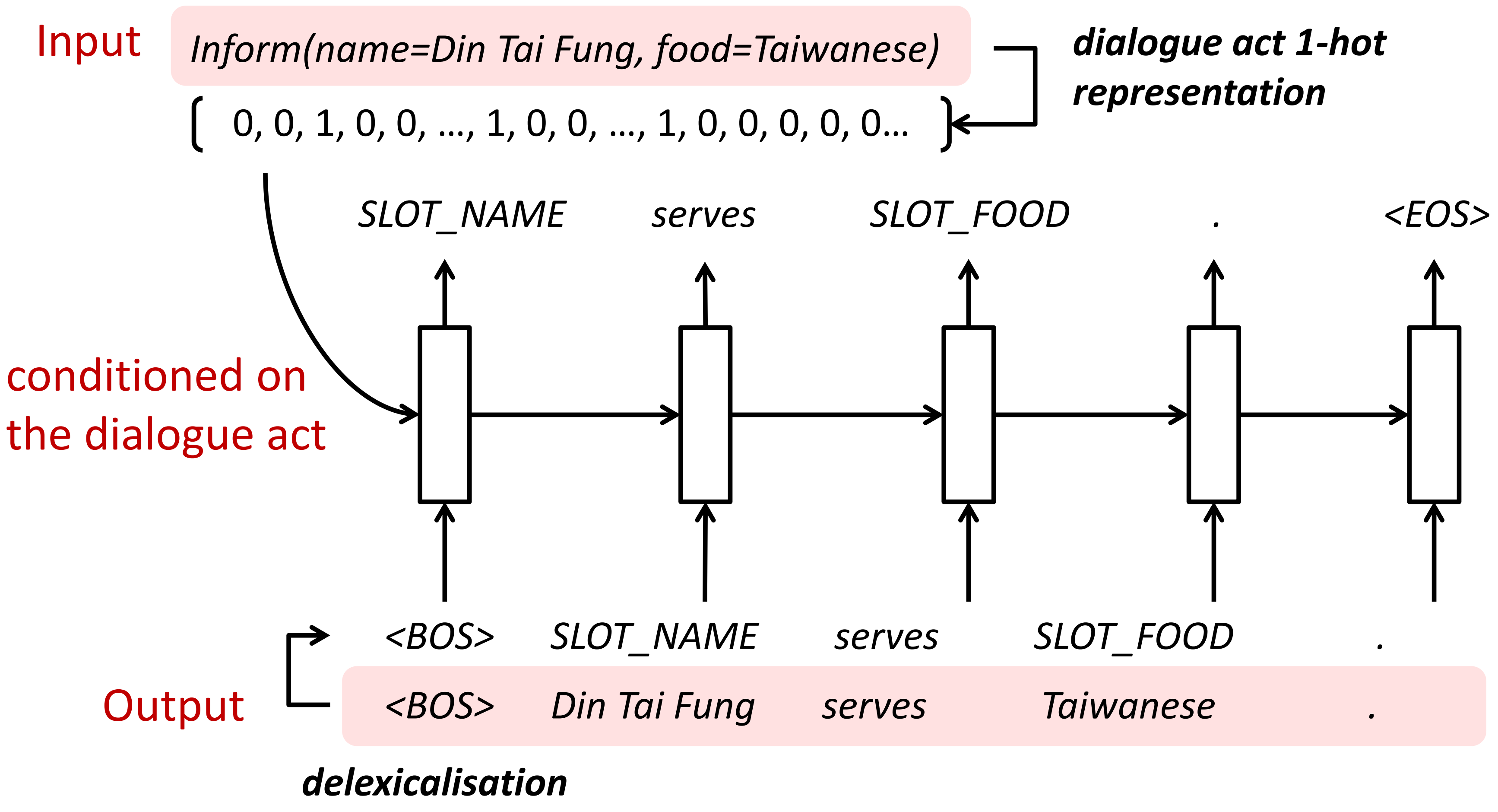
Template-Based NLG

- Define <u>a set of rules</u> to map frames to natural language

Semantic Frame	Natural Language
confirm()	"Please tell me more about the product you are looking for."
confirm(area=$V)	"Do you want somewhere in the $V?"
confirm(food=$V)	"Do you want a $V restaurant?"
confirm(food=$V,area=$W)	"Do you want a $V restaurant in the $W."

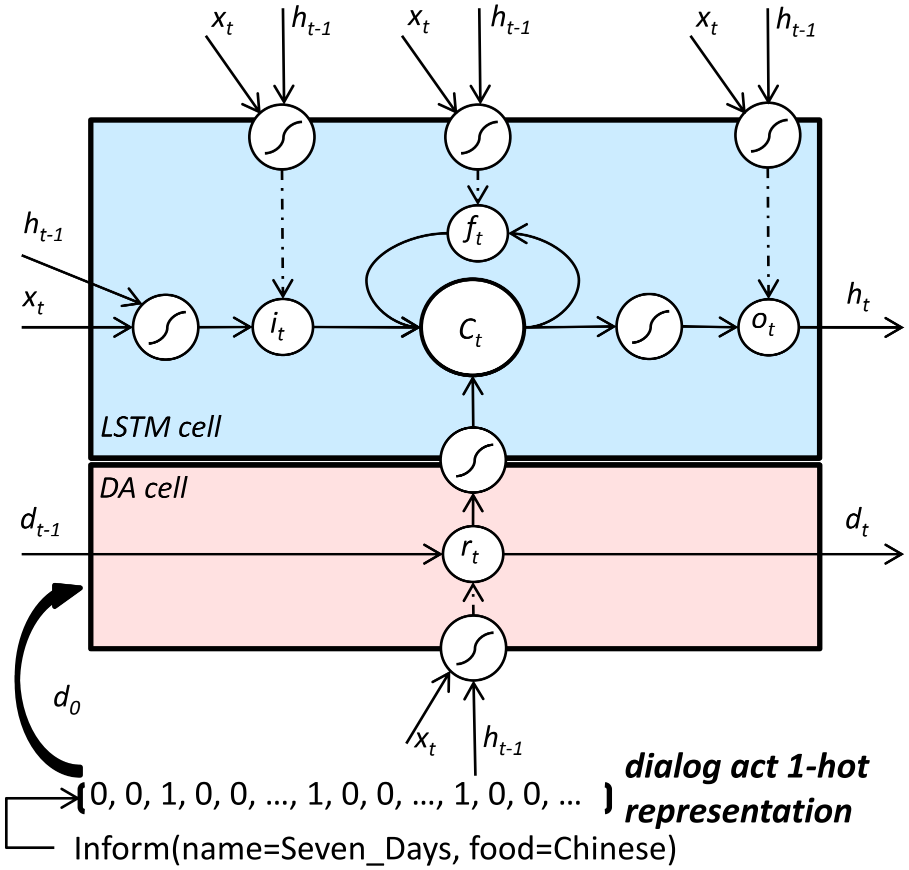
Pros: simple, error-free, easy to control
Cons: time-consuming, rigid, poor scalability

RNN-Based LM NLG (Wen et al., 2015)



Semantic Conditioned LSTM (Wen et al., 2015)

- Issue: semantic repetition
  - Din Tai Fung is a great Taiwanese restaurant that serves Taiwanese.
  - Din Tai Fung is a child friendly restaurant, and also allows kids.



Idea: using gate mechanism to control the generated semantics (dialogue act/slots)

Structural NLG (Sharma+, 2017; Nayak+, 2017)

- Delexicalized slots do not consider the word level information

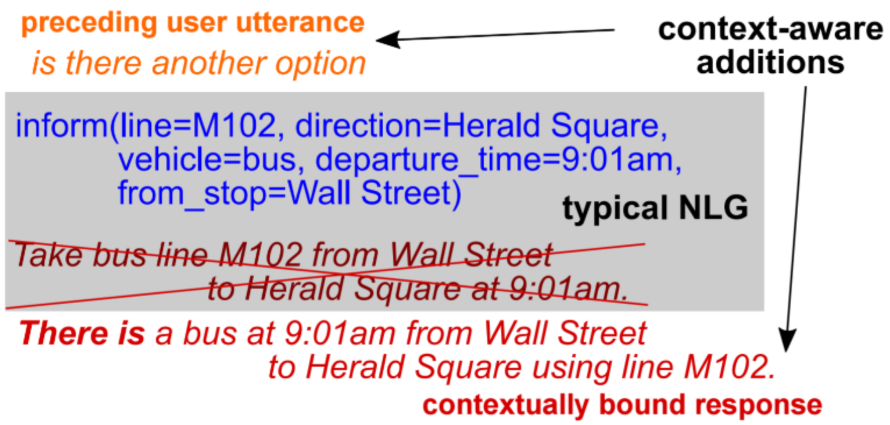
Generated output:	There are no restaurants around which serve INFORM-FOOD food.			
Delexicalized slot input:	INFORM-FOOD	✓	INFORM-FOOD	✗
Lexicalized value input:	chinese		pizza	

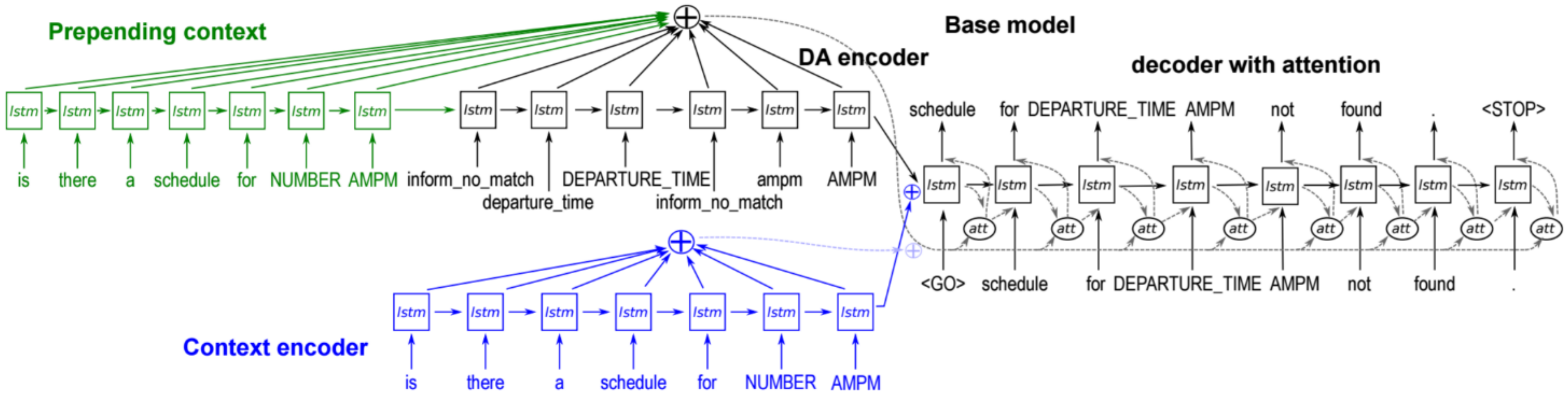
- Slot value-informed sequence to sequence models

Mention rep.	Input sequence					
SEQ	x_i	x_{i+1}	x_{i+2}	x_{i+3}	x_{i+4}	...
	decor	*decent*	service	good	cuisine	...
JOINT	x_i		x_{i+1}		x_{i+2}	
	⟨ decor, decent ⟩		⟨ service, good ⟩		⟨ cuisine, null ⟩	
CONCAT	$x_{i,1}$	$x_{i,2}$	$x_{i+1,1}$	$x_{i+1,2}$	$x_{i+2,1}$	$x_{i+2,2}$
	decor	decent	service	good	cuisine	null

Contextual NLG (Dušek and Jurčíček, 2016)

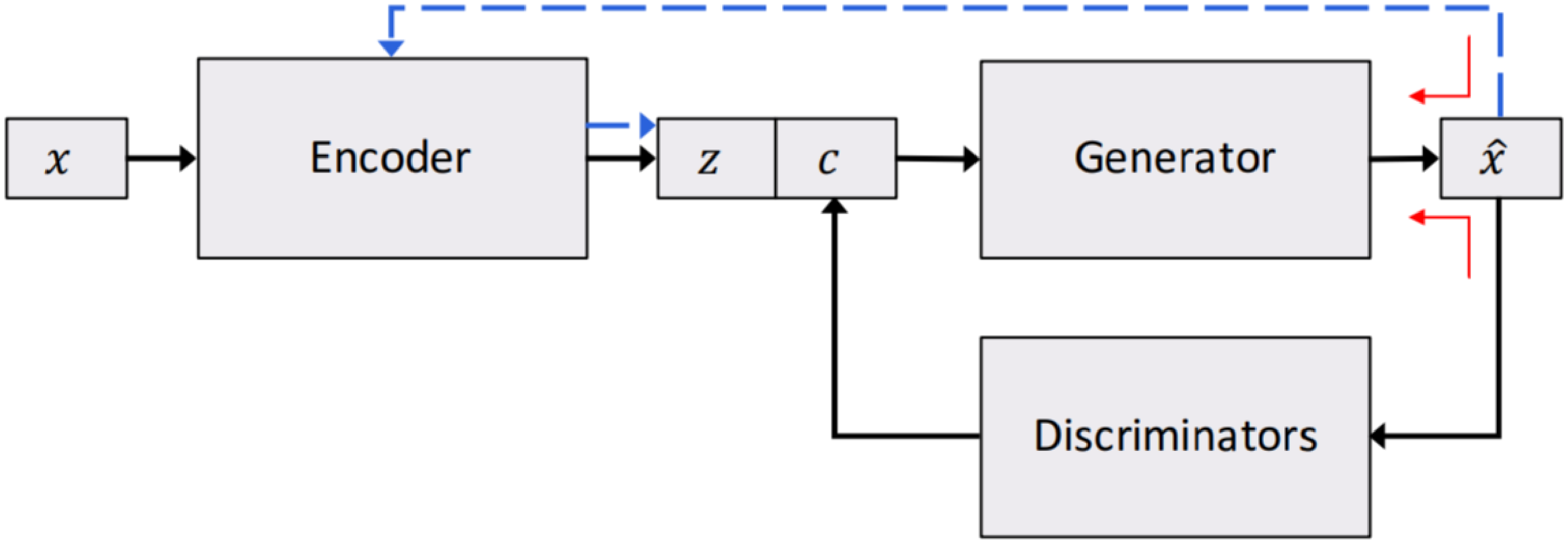
- Goal: adapting users' speaking way, providing context-aware responses
  - Context encoder
  - Seq2Seq model





Controlled Text Generation (Hu et al., 2017)

- Idea: NLG based on generative adversarial network (GAN) framework
  - *c*: targeted sentence attributes

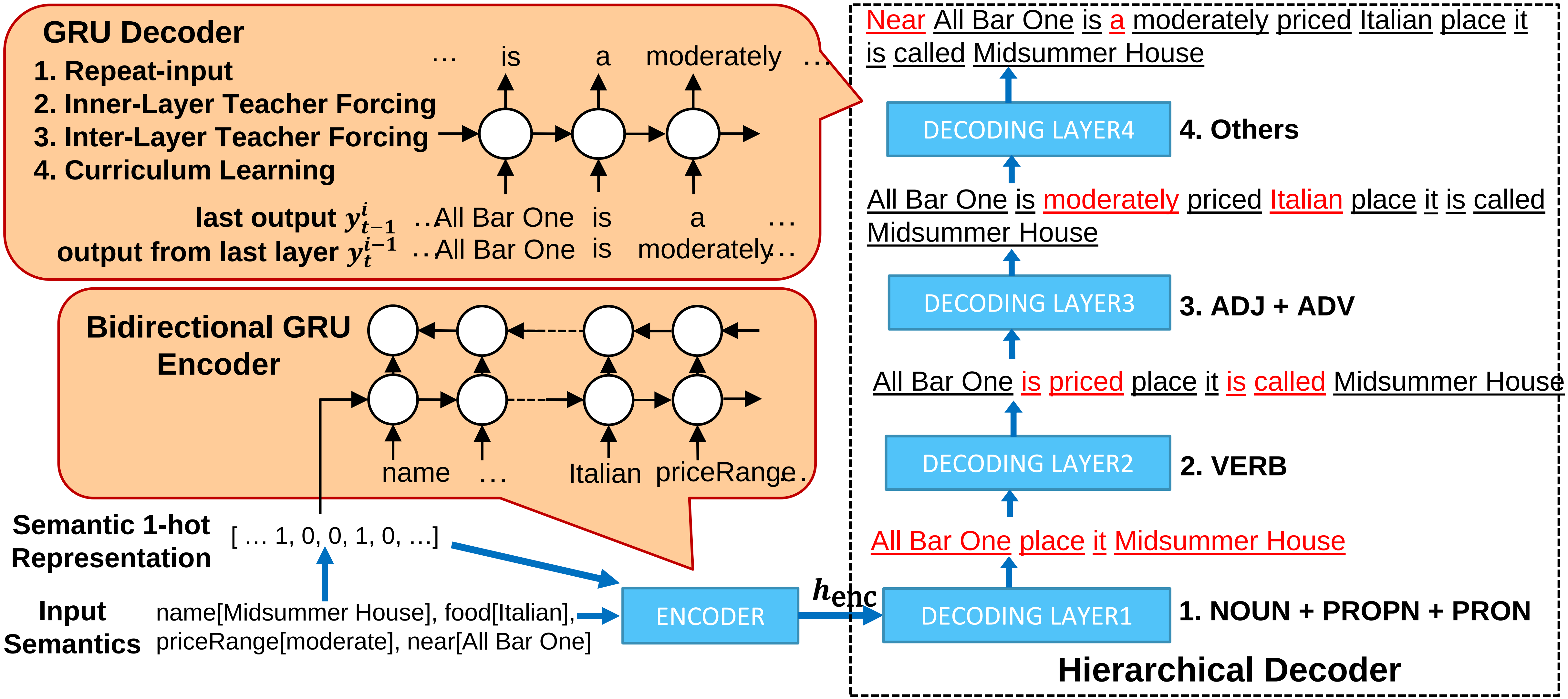


Issues in NLG

- Issue
  - NLG tends to generate shorter sentences
  - NLG may generate grammatically-incorrect sentences
- Solution
  - Generate word patterns in an order
  - Consider linguistic patterns

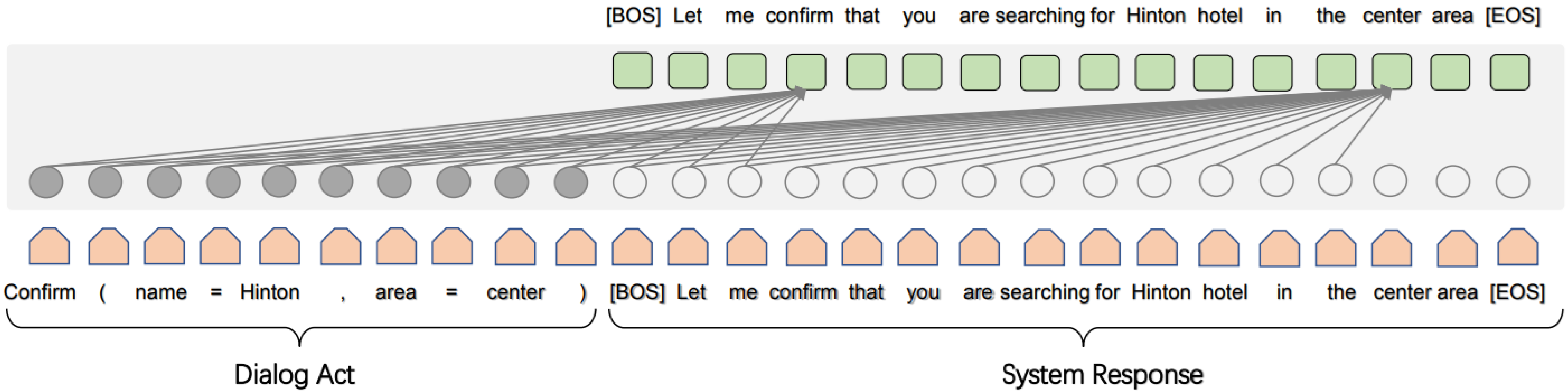
Hierarchical NLG w/ Linguistic Patterns

(Su et al., 2018)



Fine-Tuning Pre-Trained GPT-2

- Fine-tuning for conditional generation



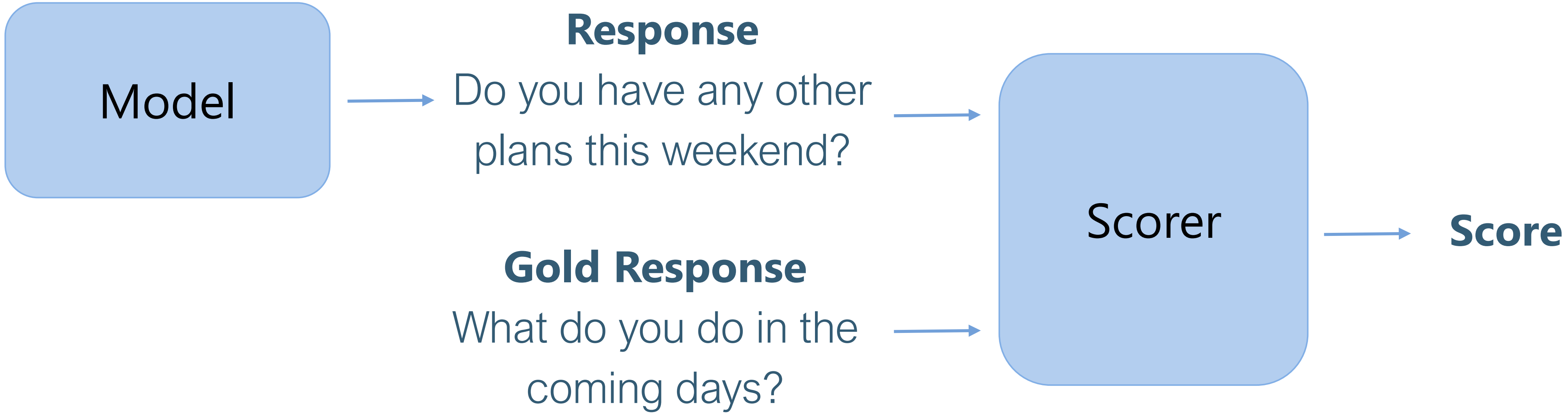
Pre-trained models have better capability of generating fluent sentences

NLG Evaluation

System Action
inform(name=B&B) ⇨ **System Response**
I have book a hotel B&B for you.

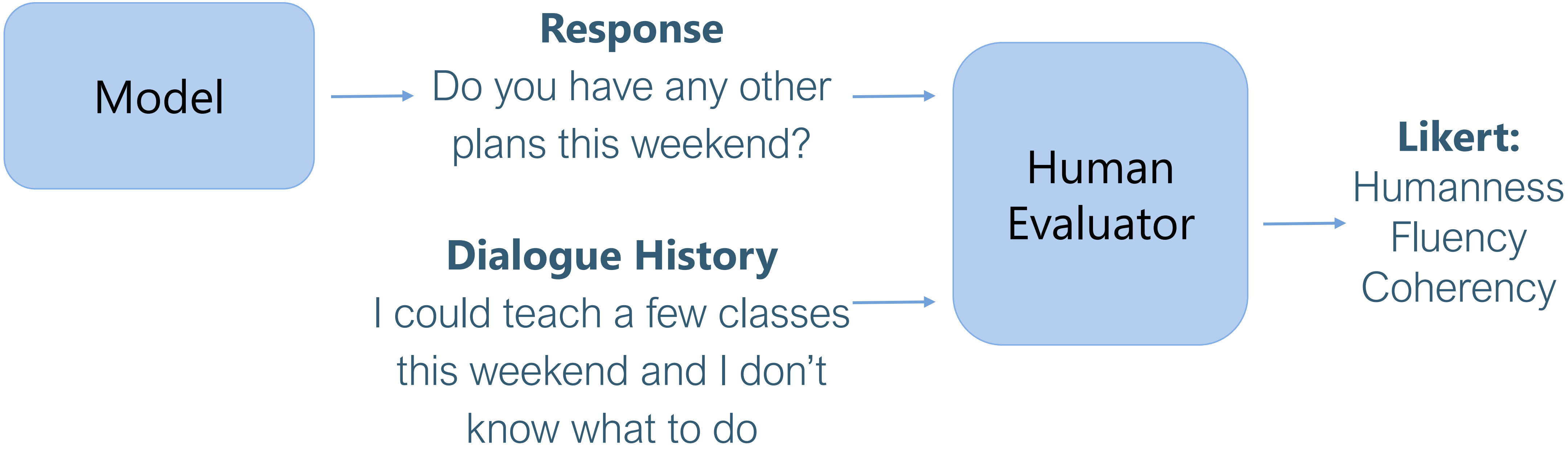
- Automatic metrics
- Human evaluation

Automatic Evaluation



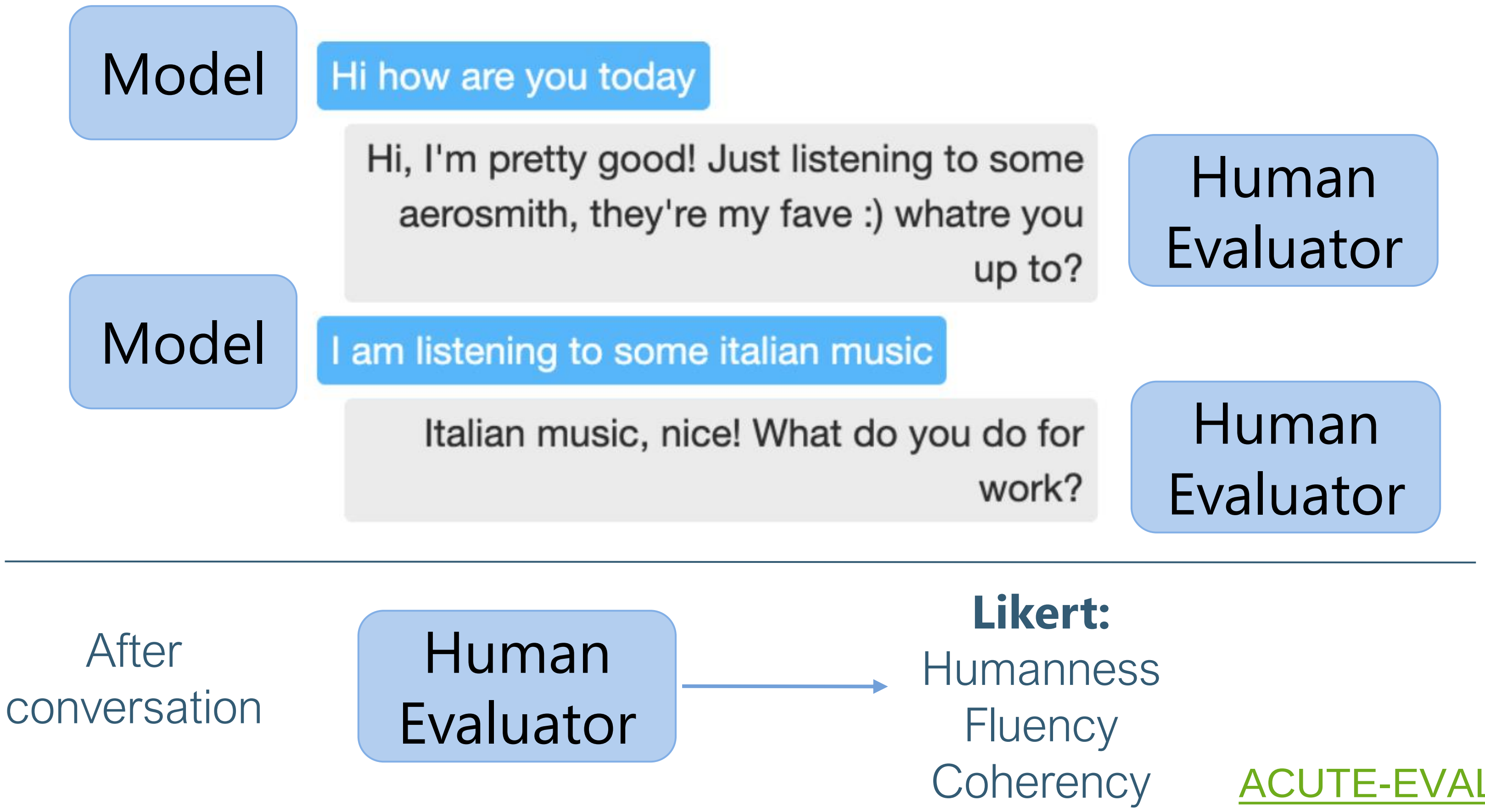
- Perplexity ⇒ how likely the model is to generate the gold response
- N-gram overlapping ⇒ BLEU etc.
- Slot error rate ⇒ whether the given slots are mentioned
- Distinct N-grams ⇒ response diversity

Human Evaluation Likert



- Judges are asked to give ratings 0-5 according to "Humanness, Fluency and Coherence"

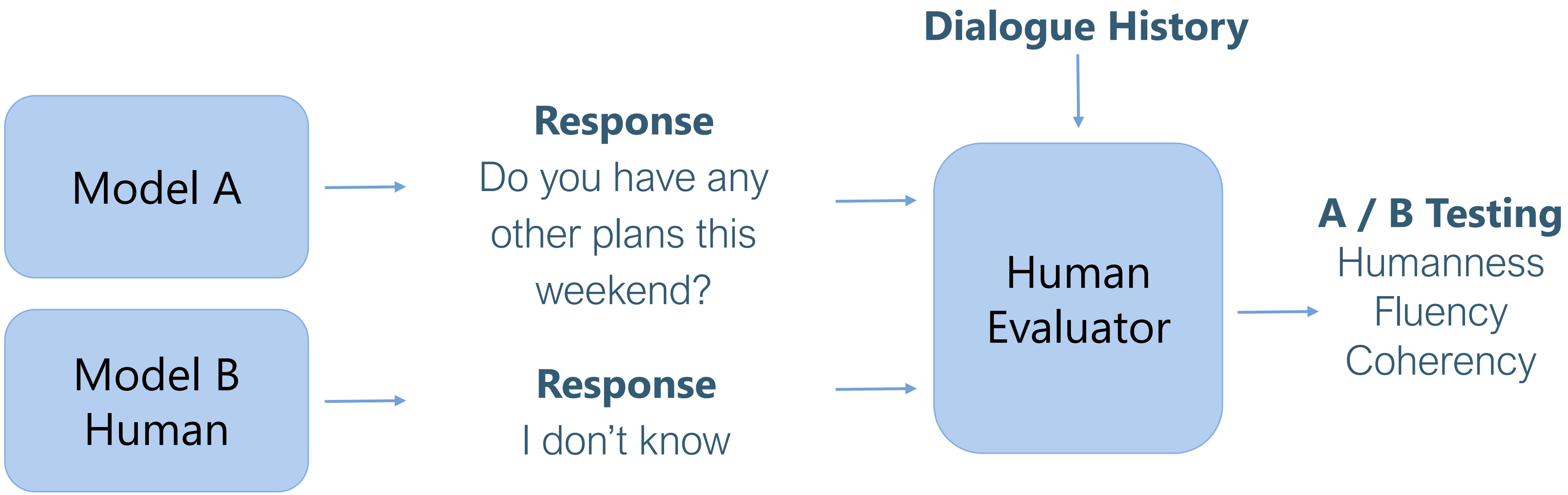
Human Evaluation Dynamic Likert



ACUTE-EVAL (Li et.al. 2019)

- Human judge interacts with the model and give ratings 0-5 according to “Humanness, Fluency and Coherence”

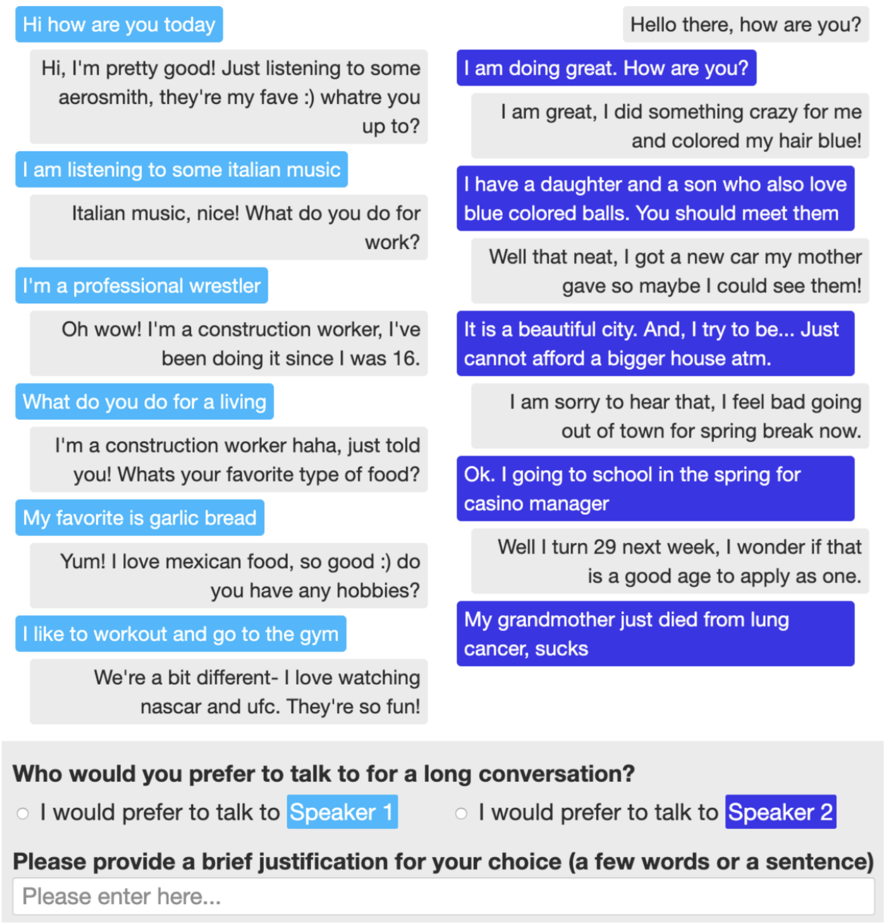
Human Evaluation A/B



- Judges are asked to choose the best one according to "Humanness, Fluency and Coherence"

Human Evaluation A/B Dynamic

- Judges interacts with two systems and chooses which is better according to "Humanness, Fluency and Coherence"
  ➔ dialogue-level evaluation



ACUTE-EVAL (Li et.al. 2019)

Conclusion Remarks