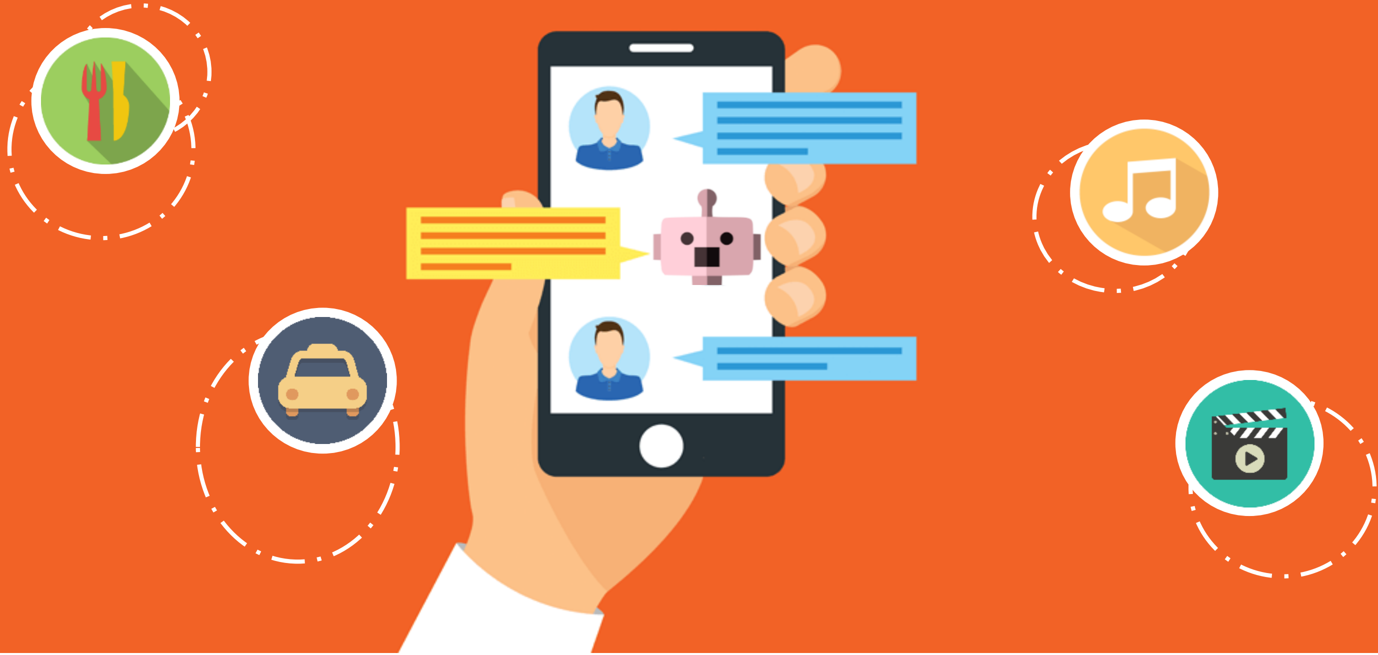
Yun-Nung (Vivian) Chen
國立臺灣大學 National Taiwan University


Asli Celikyilmaz
Microsoft Research


Dilek Hakkani-Tür
amazon alexa

deepdialogue.miulab.tw

Deep Learning for Dialogue Systems

Outline

- Introduction & Background
  - Neural Networks
  - Reinforcement Learning
- Modular Dialogue System
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management (DM)
    - Dialogue State Tracking (DST)
    - Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - End-to-End Neural Dialogue Systems
- Evaluation
- Recent Trends on Learning Dialogues

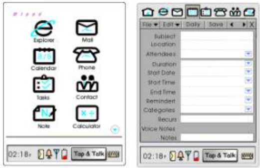
Introduction & Background

Neural Networks

Reinforcement Learning

Brief History of Dialogue Systems

Multi-modal systems
e.g., Microsoft MiPad, Pocket PC



TV Voice Search
e.g., Bing on Xbox

Virtual Personal Assistants

Task-specific argument extraction
(e.g., Nuance, SpeechWorks)
User: "I want to fly from Boston to New York next week."

2017

Early 2000s

IBM WATSON

Early 1990s

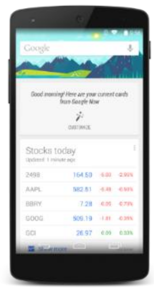
Intent Determination
(Nuance's Emily™, AT&T HMIHY)
User: "Uh…we want to move…we want to change our phone line from this house to another house"

DARPA
CALO Project

Keyword Spotting
(e.g., AT&T)
System: "Please say collect, calling card, person, third number, or operator"

Language Empowering Intelligent Assistant

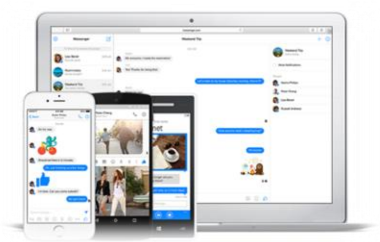
Apple Siri (2011)



Google Now (2012)
Google Assistant (2016)

Microsoft Cortana (2014)

Amazon Alexa/Echo (2014)



Facebook M & Bot (2015)

Google Home (2016)

Apple HomePod (2017)

Conversational Agents

Chit-Chat

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

Task-Oriented

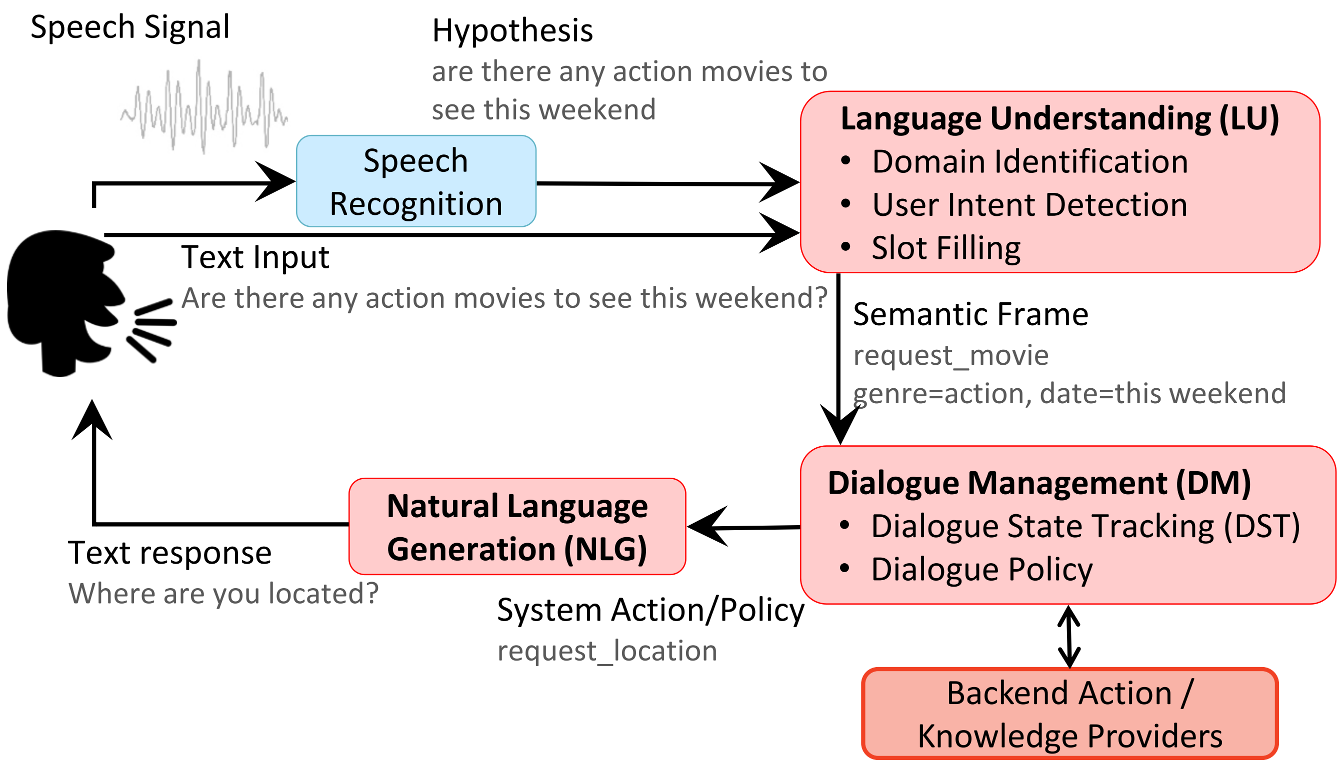
Single-domain, system-initiative → Multi-domain, contextual, mixed-initiative → End-to-end learning, massively multi-domain

Challenges

- Variability in natural language
- Robustness
- Recall/Precision Trade-off
- Meaning Representation
- Common Sense, World Knowledge
- Ability to learn
- Transparency

Task-Oriented Dialogue System (Young, 2000)

http://rsta.royalsocietypublishing.org/content/358/1769/1389.short



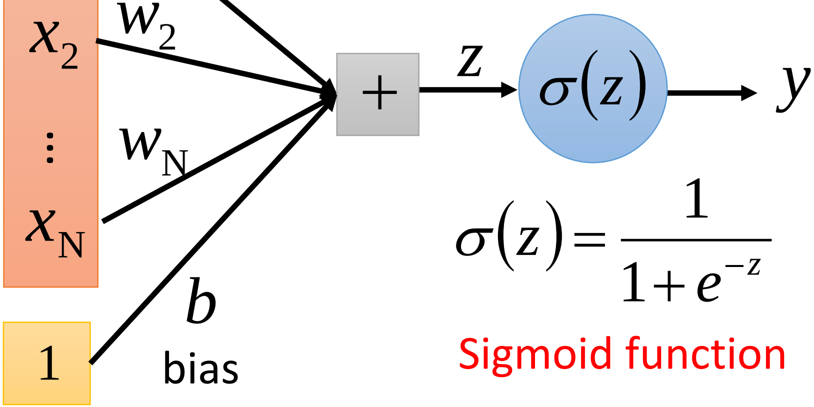
Outline

- Introduction & Background
  - **Neural Networks**
  - Reinforcement Learning
- Modular Dialogue System
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management (DM)
    - Dialogue State Tracking (DST)
    - Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - End-to-End Neural Dialogue Systems
- System Evaluation
- Recent Trends on Learning Dialogues

A Single Neuron

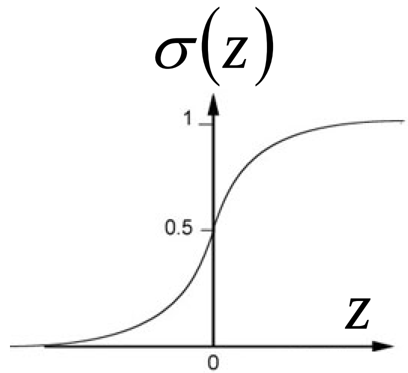
$$y = h_{w,b}(x) = \sigma(w^T x + b)$$

Activation function



$$\sigma(z) = \frac{1}{1+e^{-z}}$$

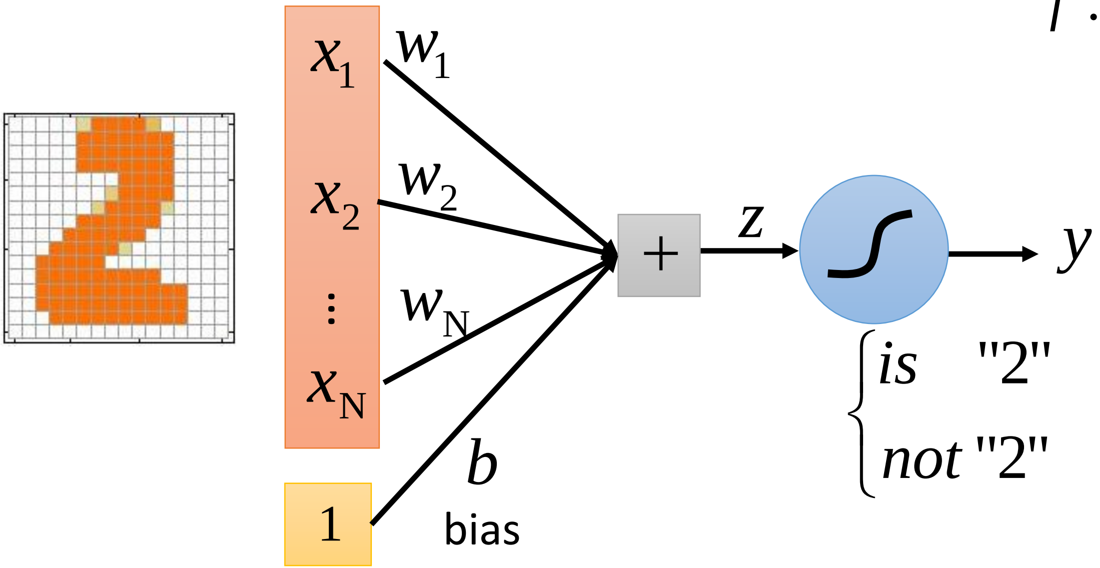
Sigmoid function



w, b are the parameters of this neuron

A Single Neuron

$$f : R^N \rightarrow R^M$$



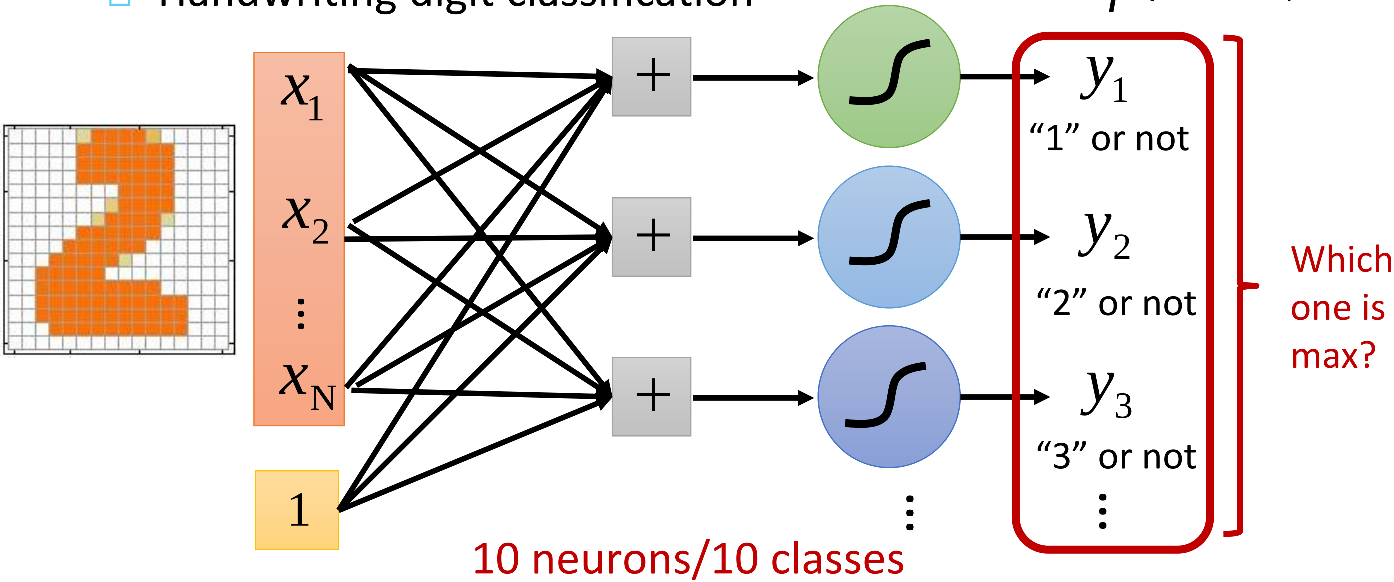
$$\begin{cases} is \quad "2" & y \geq 0.5 \\ not \ "2" & y < 0.5 \end{cases}$$

A single neuron can only handle binary classification

A Layer of Neurons

- Handwriting digit classification

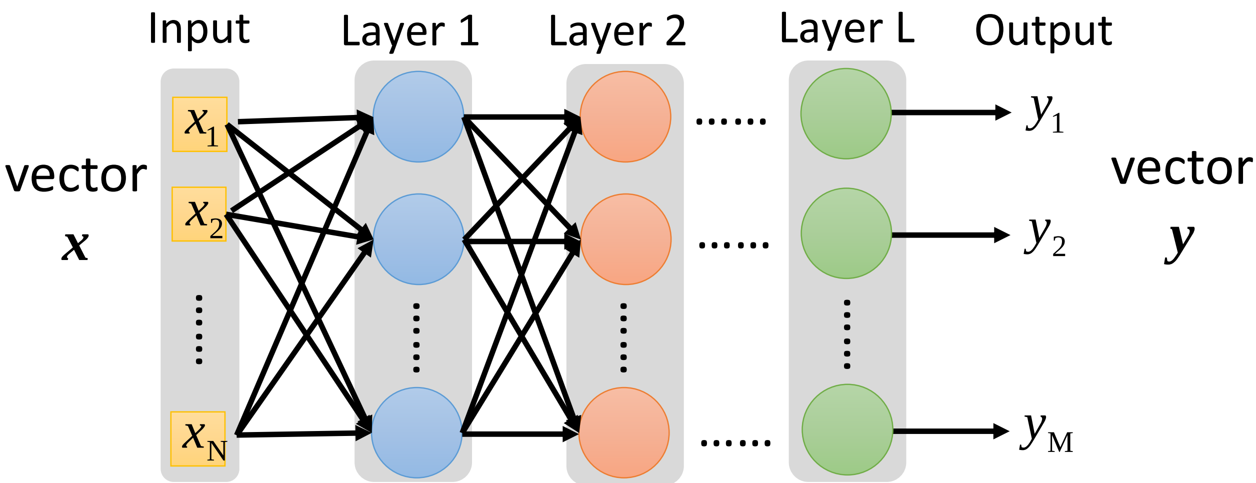
$$f : R^N \rightarrow R^M$$



A layer of neurons can handle multiple possible output, and the result depends on the max one

Deep Neural Networks (DNN)

- Fully connected feedforward network $f: R^N \rightarrow R^M$

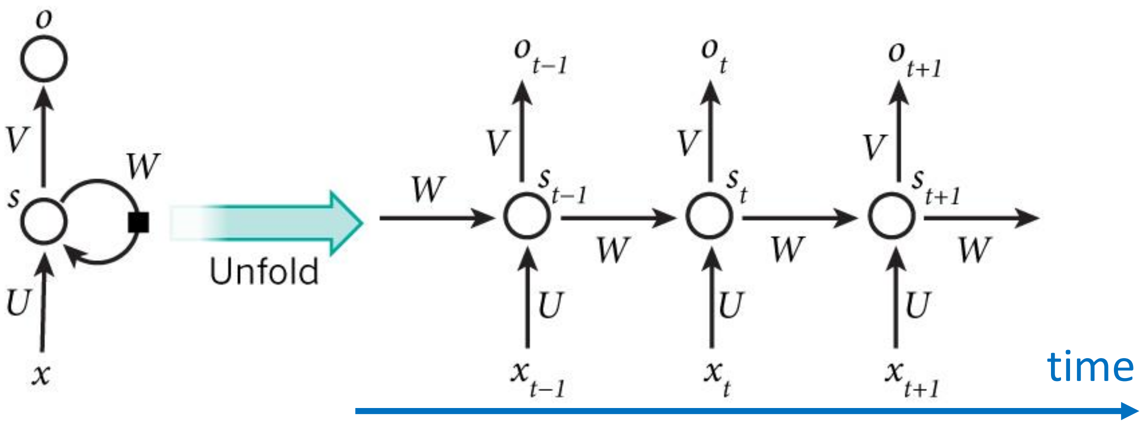


Deep NN: multiple hidden layers

Recurrent Neural Network (RNN)

$$s_t = \sigma(W s_{t-1} + U x_t)$$ $\sigma(\cdot)$: tanh, ReLU

$$o_t = \mathrm{softmax}(V s_t)$$



RNN can learn accumulated sequential information (time-series)

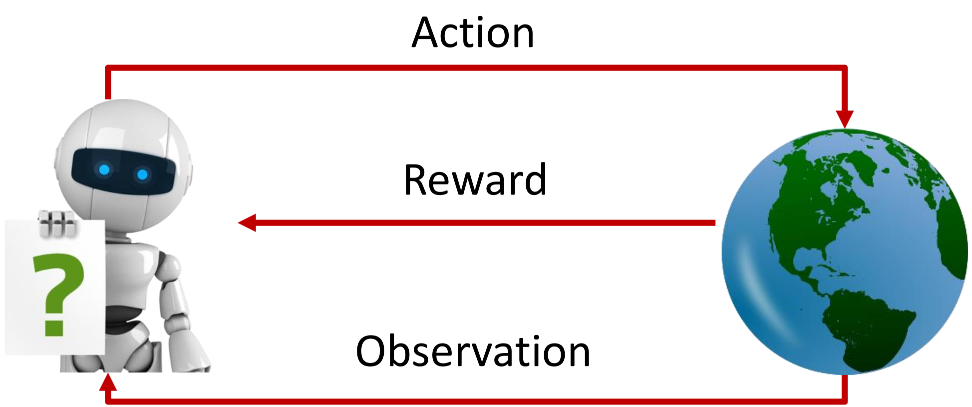
http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/

Outline

- Introduction & Background
  - Neural Networks
  - **Reinforcement Learning**
- Modular Dialogue System
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management (DM)
    - Dialogue State Tracking (DST)
    - Dialogue Policy Optimization
  - Natural Language Generation (NLG)
- System Evaluation
- Recent Trends on Learning Dialogues

Reinforcement Learning

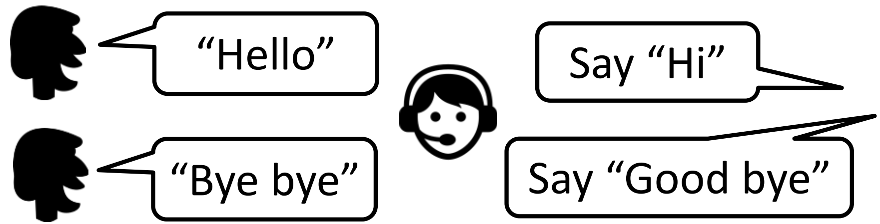
- RL is a general purpose framework for **decision making**
  - RL is for an *agent* with the capacity to *act*
  - Each *action* influences the agent’s future *state*
  - Success is measured by a scalar *reward* signal
  - Goal: *select actions to maximize future reward*



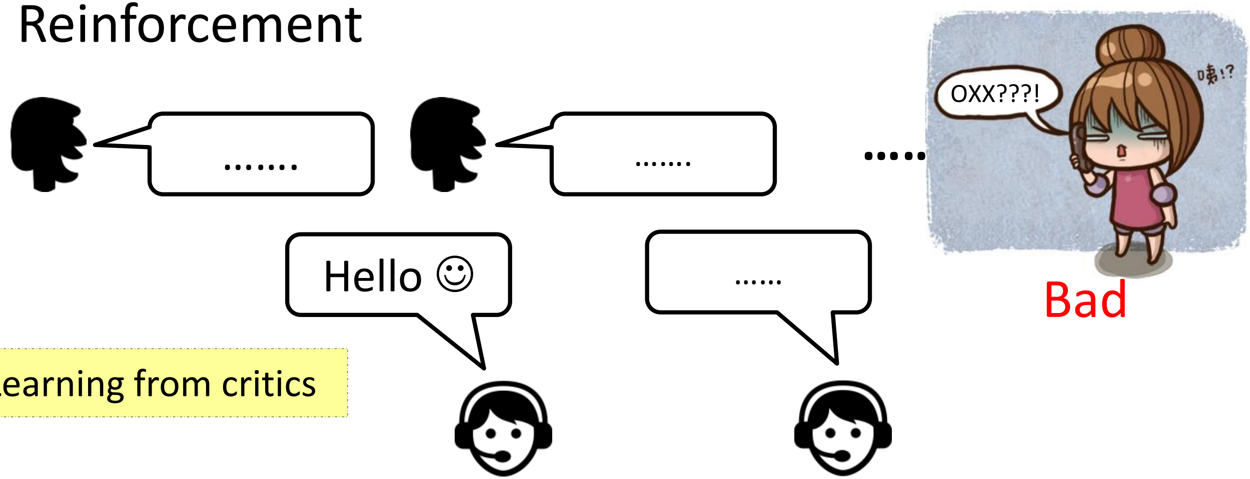
Supervised v.s. Reinforcement

- Supervised

Learning from teacher

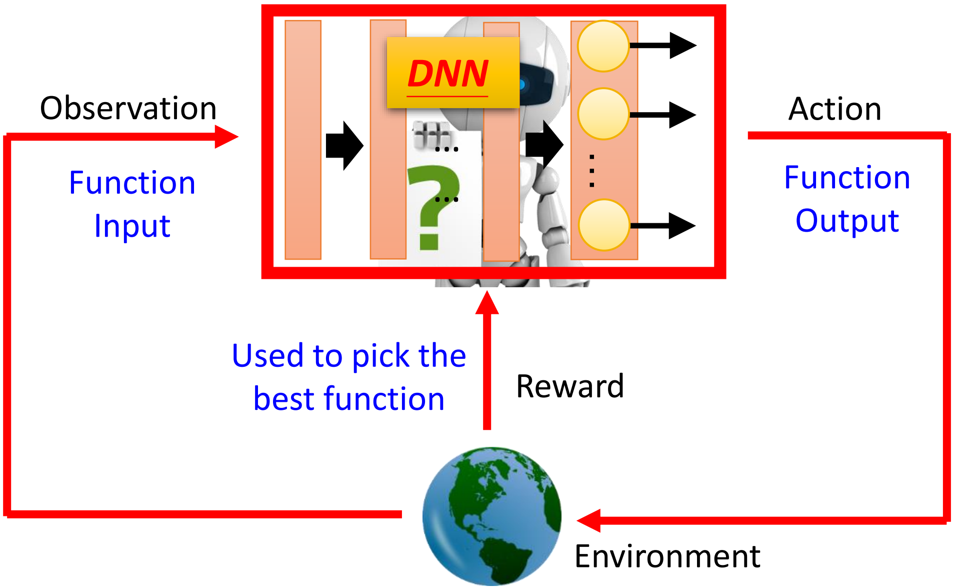


- Reinforcement



Learning from critics

Deep Reinforcement Learning



Goal: select actions that maximize the expected total reward

$$\mathbb{E}[R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots]$$

Part II

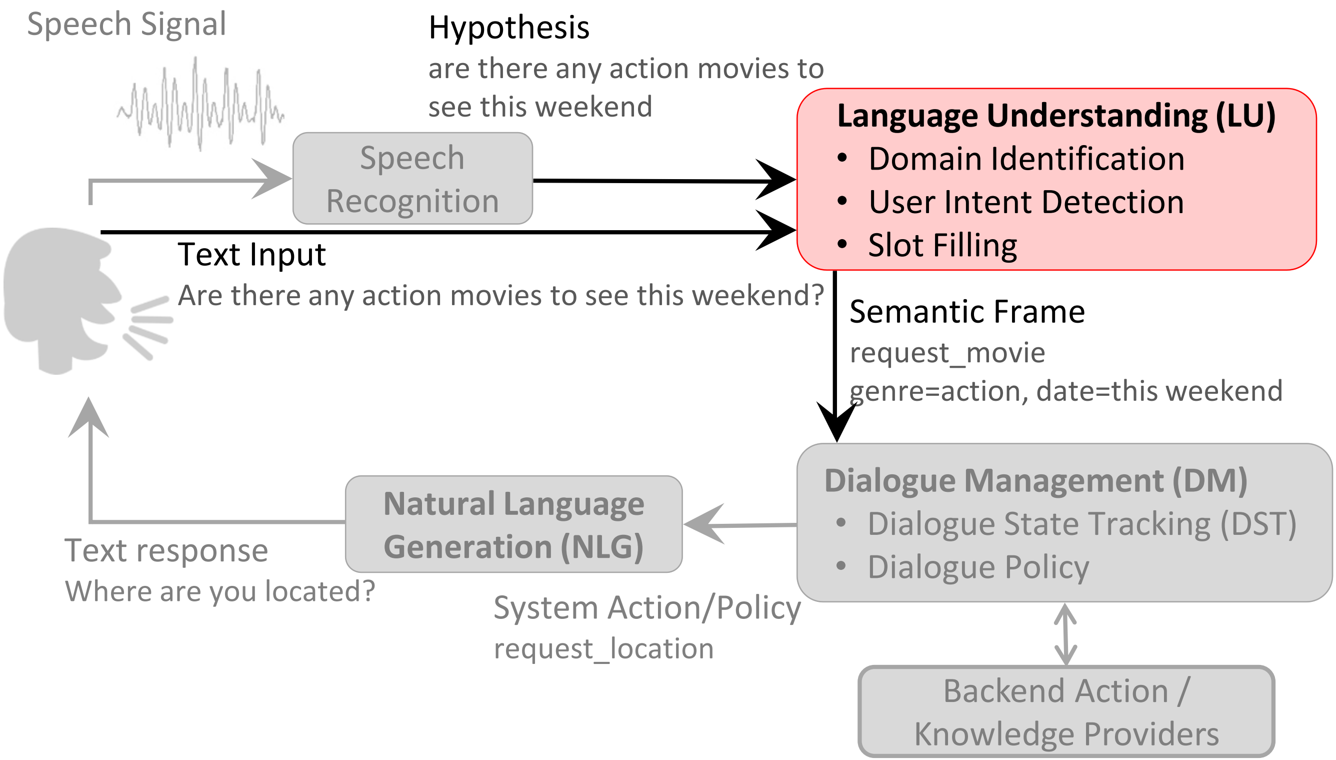
Modular Dialogue System

Outline

- Introduction & Background
  - Neural Networks
  - Reinforcement Learning
- Modular Dialogue System

  - **Spoken/Natural Language Understanding (SLU/NLU)**
  - Dialogue Management (DM)
    - Dialogue State Tracking (DST)
    - Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - End-to-End Neural Dialogue Systems
- System Evaluation
- Recent Trends on Learning Dialogues

Task-Oriented Dialogue System (Young, 2000)

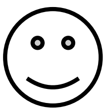


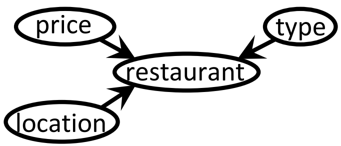
Semantic Frame Representation

- Requires a domain ontology: early connection to backend
- Contains **core content (intent, a set of slots with fillers)**

Restaurant Domain

find me a cheap taiwanese restaurant in oakland

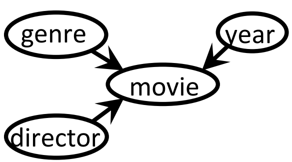




find_restaurant (price="cheap", type="taiwanese", location="oakland")

Movie Domain

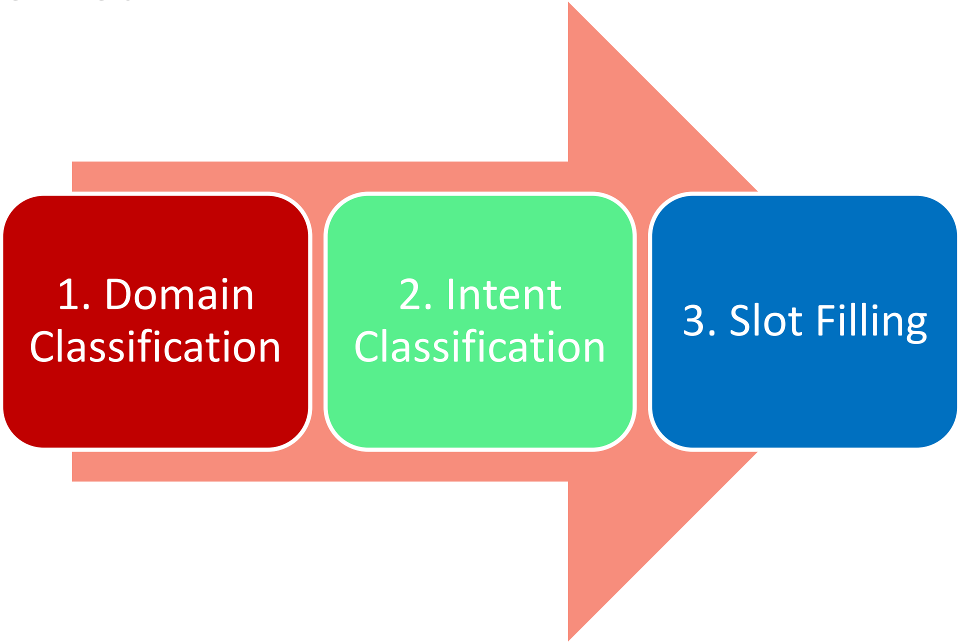
show me action movies directed by james cameron



find_movie (genre="action", director="james cameron")

Language Understanding (LU)

- Pipelined



LU – Domain/Intent Classification

As an **utterance classification** task

- Given a collection of utterances u_i with labels c_i, $D = \{(u_1,c_1),\dots,(u_n,c_n)\}$ where $c_i \in C$, train a model to estimate labels for new utterances u_k.

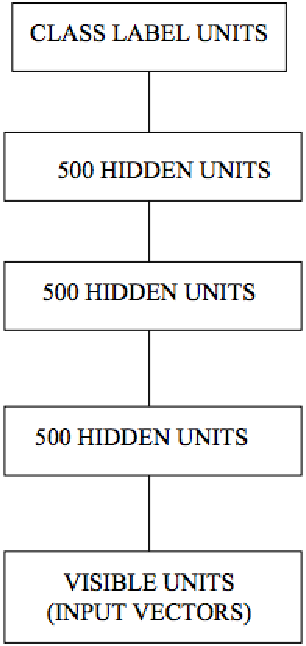
find me a cheap taiwanese restaurant in oakland

Domain	Intent
Movies	find_movie, buy_tickets
Restaurants	find_restaurant, find_price, book_table
Music	find_lyrics, find_singer
Sports	…
…	

Domain/Intent Classification (Sarikaya+, 2011)

http://ieeexplore.ieee.org/abstract/document/5947649/

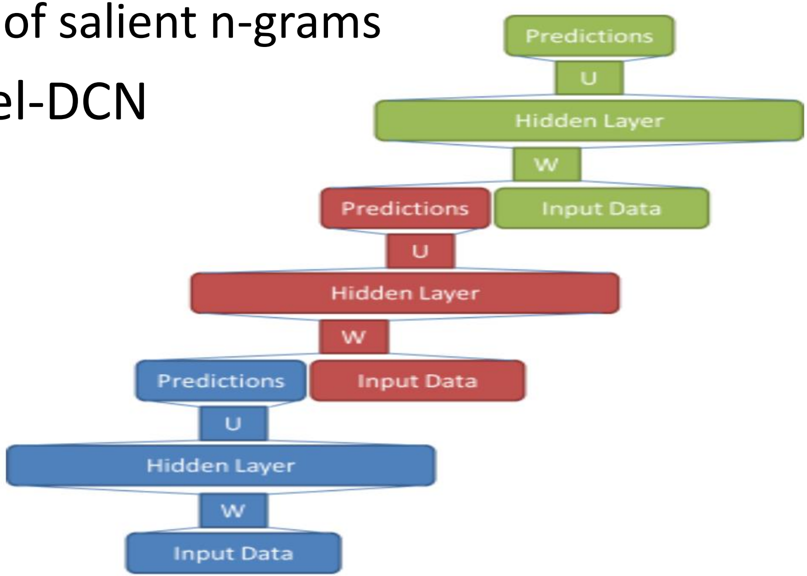
- Deep belief nets (DBN)
  - Unsupervised training of weights
  - Fine-tuning by back-propagation
  - Compared to MaxEnt, SVM, and boosting



Domain/Intent Classification (Tur+, 2012; Deng+, 2012)

http://ieeexplore.ieee.org/abstract/document/6289054/; http://ieeexplore.ieee.org/abstract/document/6424224/

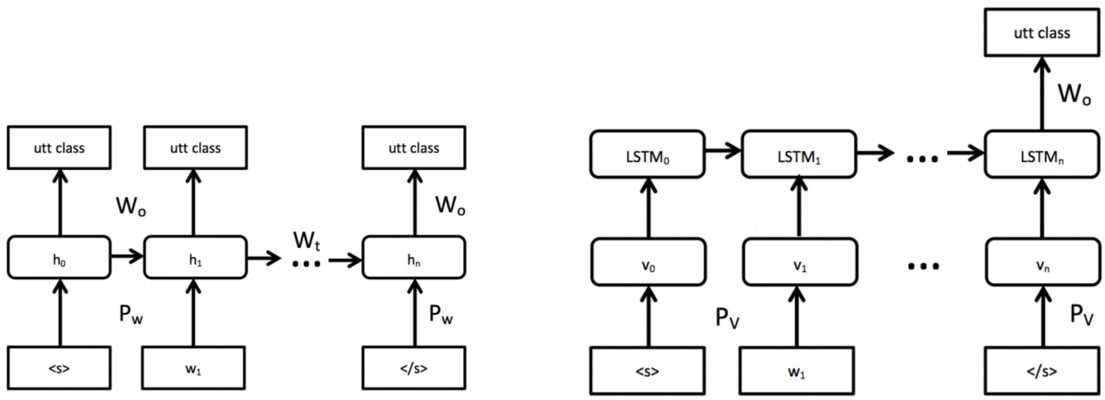
- Deep convex networks (DCN)
  - Simple classifiers are stacked to learn complex functions
  - Feature selection of salient n-grams
- Extension to kernel-DCN



Domain/Intent Classification (Ravuri & Stolcke, 2015)

https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/RNNLM_addressee.pdf

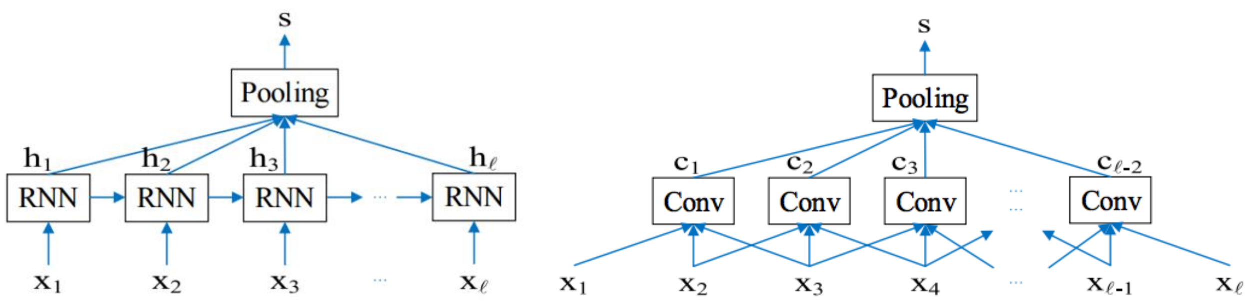
- RNN and LSTMs for utterance classification



Intent decision after reading all words performs better

Dialogue Act Classification (Lee & Dernoncourt, 2016)

- RNN and CNNs for dialogue act classification



LU – Slot Filling

As a sequence tagging task

- Given a collection tagged word sequences, $S=\{((w_{1,1},w_{1,2}, \ldots, w_{1,n1}), (t_{1,1},t_{1,2}, \ldots,t_{1,n1})), ((w_{2,1},w_{2,2}, \ldots,w_{2,n2}), (t_{2,1},t_{2,2}, \ldots,t_{2,n2})) \ldots\}$ where $t_i \in M$, the goal is to estimate tags for a new word sequence.

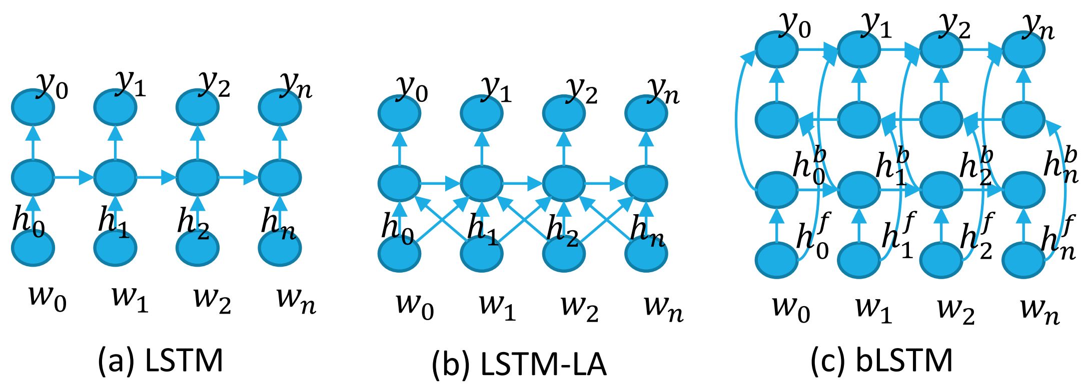
flights from Boston to New York today

	flights	from	Boston	to	New	York	today
Entity Tag	O	O	B-city	O	B-city	I-city	O
Slot Tag	O	O	B-dept	O	B-arrival	I-arrival	B-date

Slot Tagging (Yao+, 2013; Mesnil+, 2015)

http://131.107.65.14/en-us/um/people/gzweig/Pubs/Interspeech2013RNNLU.pdf; http://dl.acm.org/citation.cfm?id=2876380

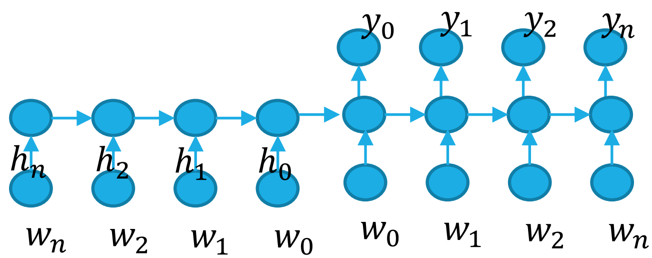
- Variations:
  a. RNNs with LSTM cells
  b. Input, sliding window of n-grams
  c. Bi-directional LSTMs



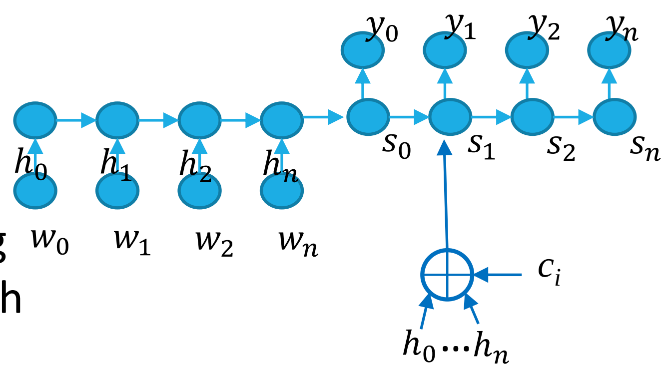
Slot Tagging (Kurata+, 2016; Simonnet+, 2015)

http://www.aclweb.org/anthology/D16-1223

- Encoder-decoder networks
  - Leverages sentence level information



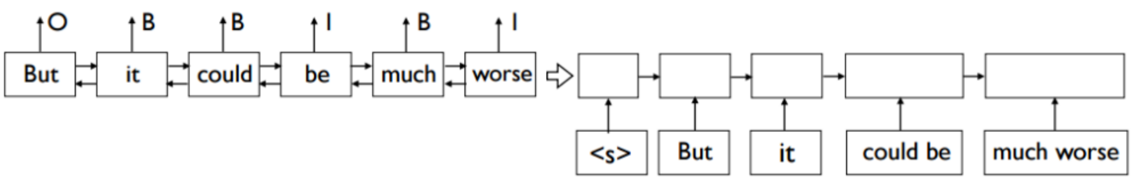
- Attention-based encoder-decoder
  - Use of attention (as in MT) in the encoder-decoder network
  - Attention is estimated using a feed-forward network with input: h_t and s_t at time t



Joint Segmentation & Slot Tagging (Zhai+, 2017)

https://arxiv.org/pdf/1701.04027.pdf

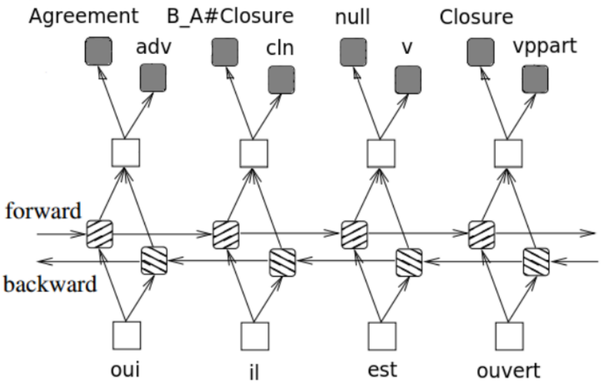
- Encoder that segments
- Decoder that tags the segments



Multi-Task Slot Tagging (Jaech+, 2016; Tafforeau+, 2016)

https://arxiv.org/abs/1604.00117; http://www.sensei-conversation.eu/wp-content/uploads/2016/11/favre_is2016b.pdf

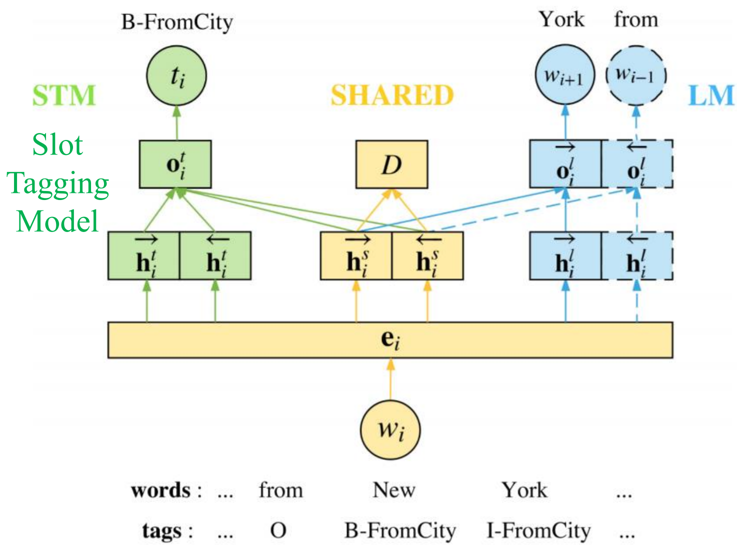
- Multi-task learning
  - Goal: exploit data from domains/tasks with a lot of data to improve ones with less data
  - Lower layers are shared across domains/tasks
  - Output layer is specific to task



Semi-Supervised Slot Tagging (Lan+, 2018)

https://speechlab.sjtu.edu.cn/papers/oyl11-lan-icassp18.pdf

Idea: language understanding objective can enhance other tasks



Algorithm 1: Adversarial Multi-task Learning for SLU

Input : Labeled training data $\{(\mathbf{w}^l, \mathbf{t}^l)\}$

Unlabeled data $\{\mathbf{w}^u\}$

Output: Adversarially enhanced slot tagging model

1 Initialize parameters $\{\theta^s, \theta^t, \theta^l, \theta^d\}$ randomly.

2 **repeat**

```
/* Sample from {(w^l, t^l)} */
```

3 Train the STM and shared model by Eq.(8).

4 Train the task discriminator and the shared model by Eq.(6) or Eq.(7) as slot tagging task ($y = 1$).

```
/* Sample from {w^l} and {w^u} */
```

5 Train the LM and shared models by Eq.(9) (and Eq.(10) for BLM).

6 Train the task discriminator and the shared model by Eq.(6) or Eq.(7) as LM task ($y = 0$).

7 **until** *convergence*;

BLM exploits the *unsupervised knowledge*, the *shared-private framework* and *adversarial training* make the slot tagging model more generalized

LU Evaluation

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

Joint Semantic Frame Parsing (Hakkani-Tur+, 2016; Liu & Lane, 2016)

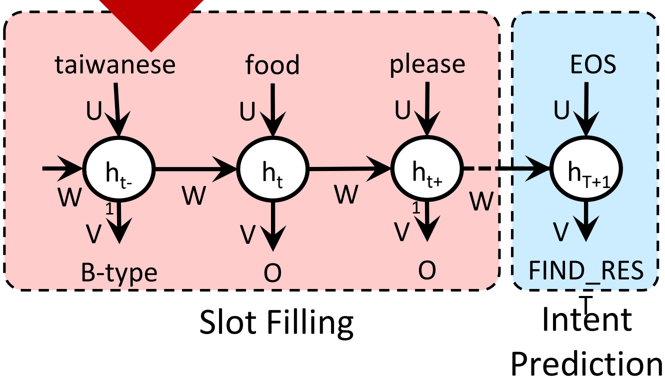
https://www.microsoft.com/en-us/research/wp-content/uploads/2016/06/IS16_MultiJoint.pdf; https://arxiv.org/abs/1609.01454

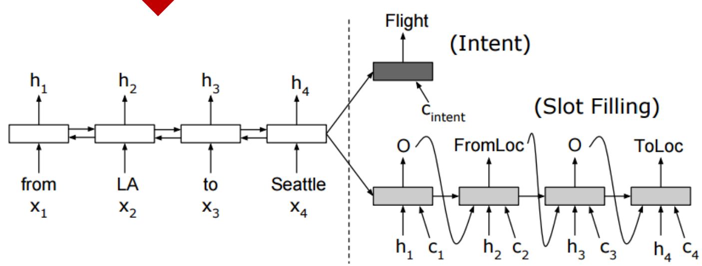
Sequence-based (Hakkani-Tur+, 2016)

- Slot filling and intent prediction in the same output sequence

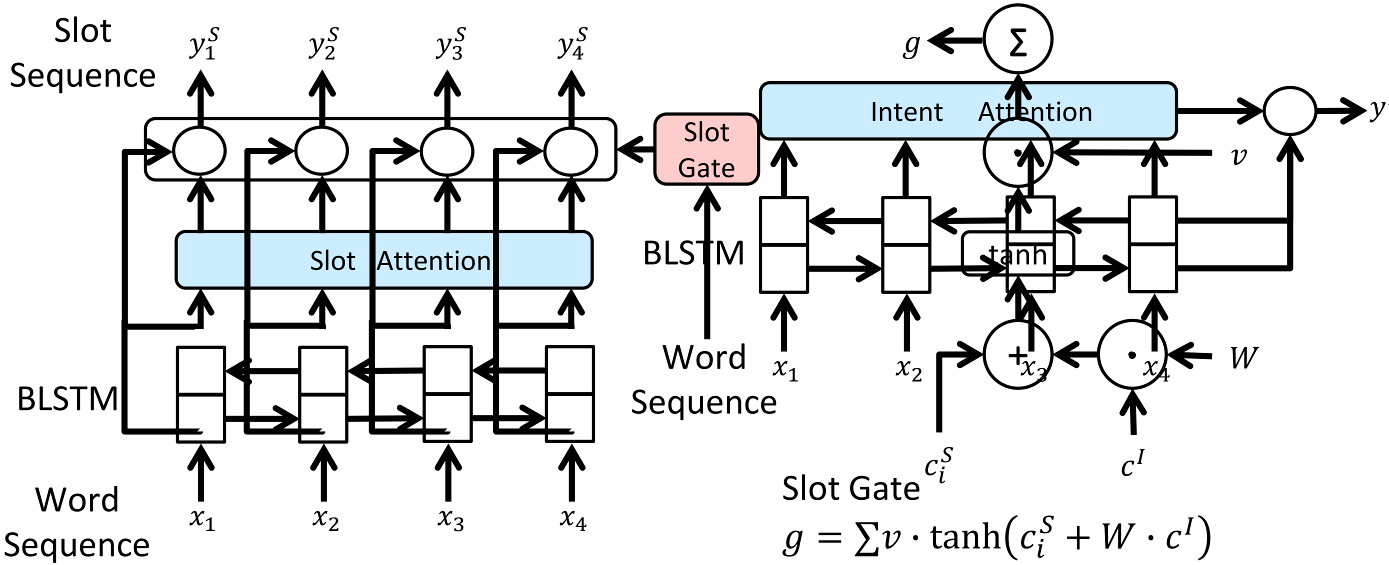
Parallel (Liu & Lane, 2016)

- Intent prediction and slot filling are performed in two branches





Slot-Gated Joint SLU (Goo+, 2018)



Slot Gate

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

Slot Prediction

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

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

Contextual LU

Domain Identification → Intent Prediction → Slot Filling

D communication *I* send_email

Single Turn

U	just	sent	email	to	bob	about	fishing	this	weekend
S	O	O	O	O	B-contact_name	O	B-subject	I-subject	I-subject

→ send_email(contact_name="bob", subject="fishing this weekend")

Multi-Turn

U_1	send	email	to	bob
S_1				B-contact_name

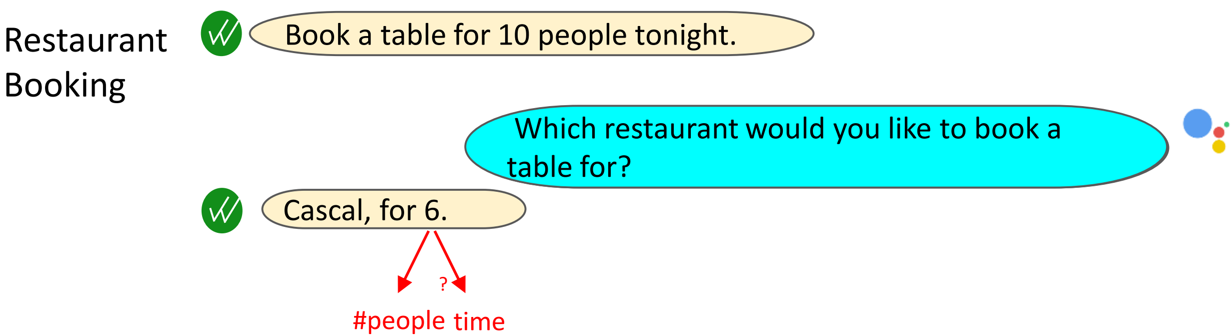
→ send_email(contact_name="bob")

U_2	are	we	going	to	fish	this	weekend
S_2	B-message	I-message	I-message	I-message	I-message	I-message	I-message

→ send_email(message="are we going to fish this weekend")

Contextual LU

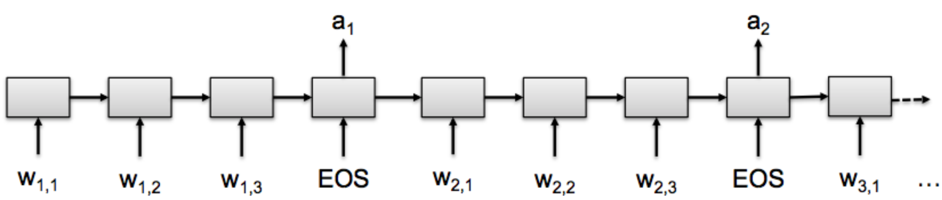
- User utterances are highly ambiguous in isolation



Contextual LU (Bhargava+, 2013; Hori+, 2015)

https://www.merl.com/publications/docs/TR2015-134.pdf

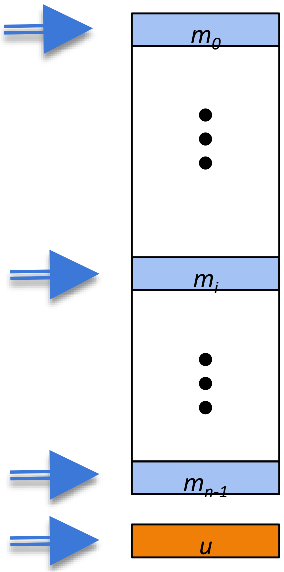
- Leveraging contexts
  - Used for individual tasks
- Seq2Seq model
  - Words are input one at a time, tags are output at the end of each utterance



- Extension: LSTM with speaker role dependent layers

End-to-End Memory Networks (Sukhbaatar+, 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)

https://www.microsoft.com/en-us/research/wp-content/uploads/2016/06/IS16_ContextualSLU.pdf

1. Sentence Encoding

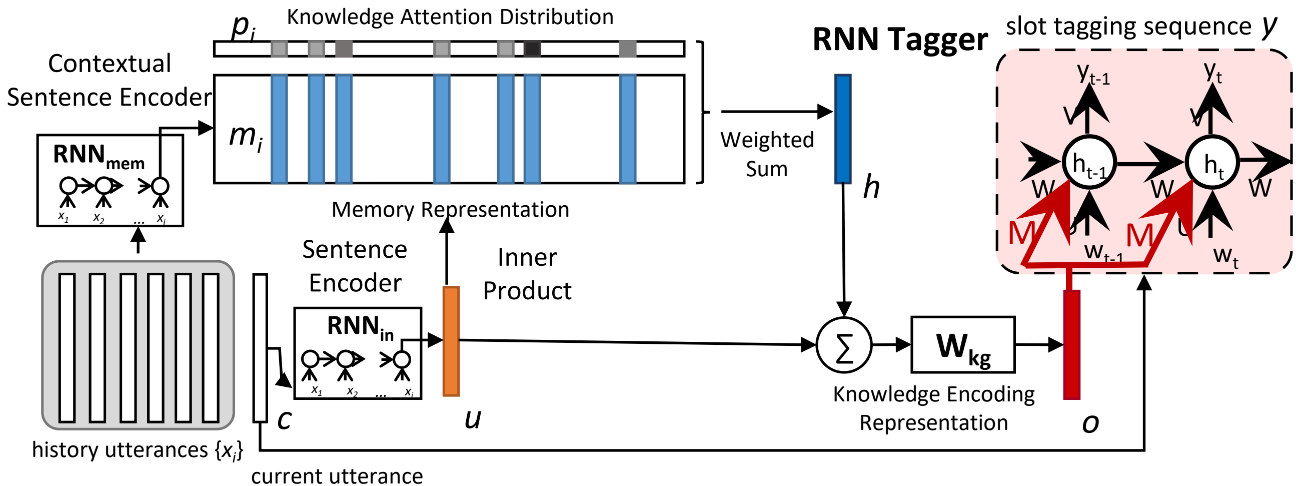
$$m_i = \mathrm{RNN}_{\mathrm{mem}}(x_i)$$
$$u = \mathrm{RNN}_{\mathrm{in}}(c)$$

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
→ track dialogue states in a latent way

Analysis of Attention

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"

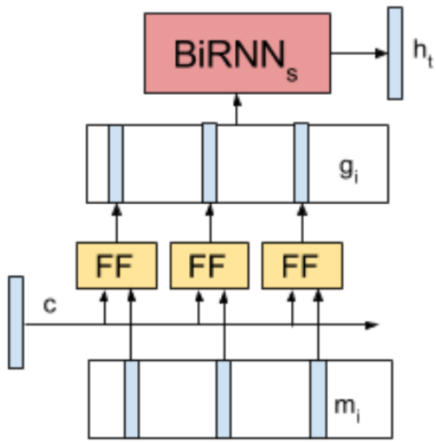
S: "okay , there is 4:10 pm , 5:40 pm and 9:20 pm" → 0.16

U: "Let's do 5:40"

Dialogue Encoder Network (Bapna+, 2017)

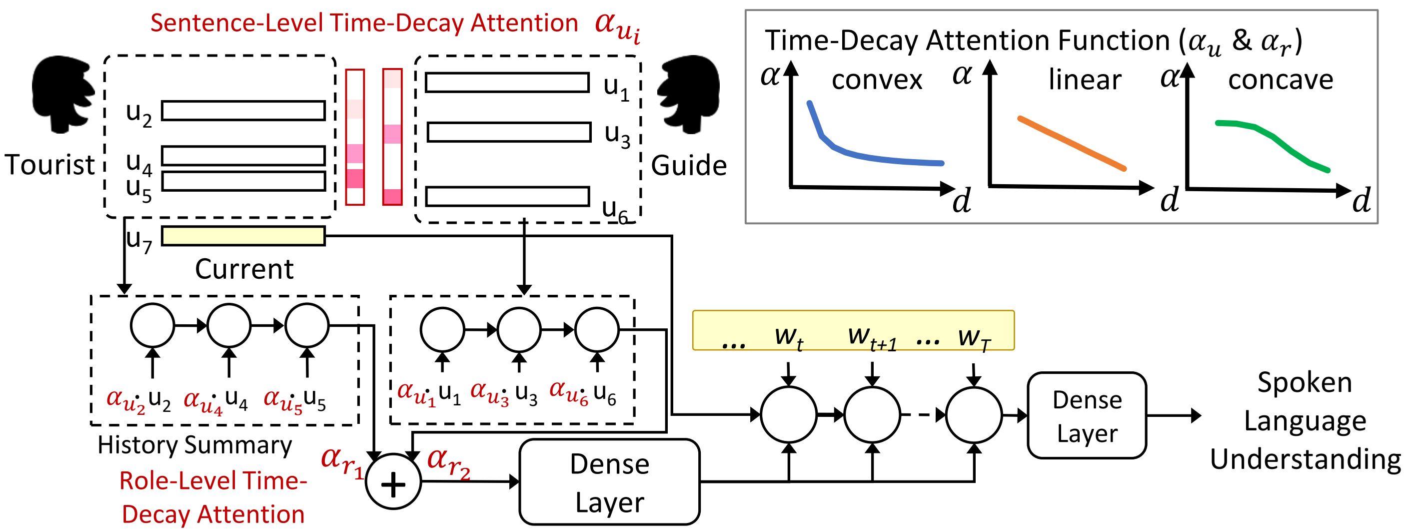
http://aclweb.org/anthology/W17-5514

- Past and current turn encodings input to a feed forward network

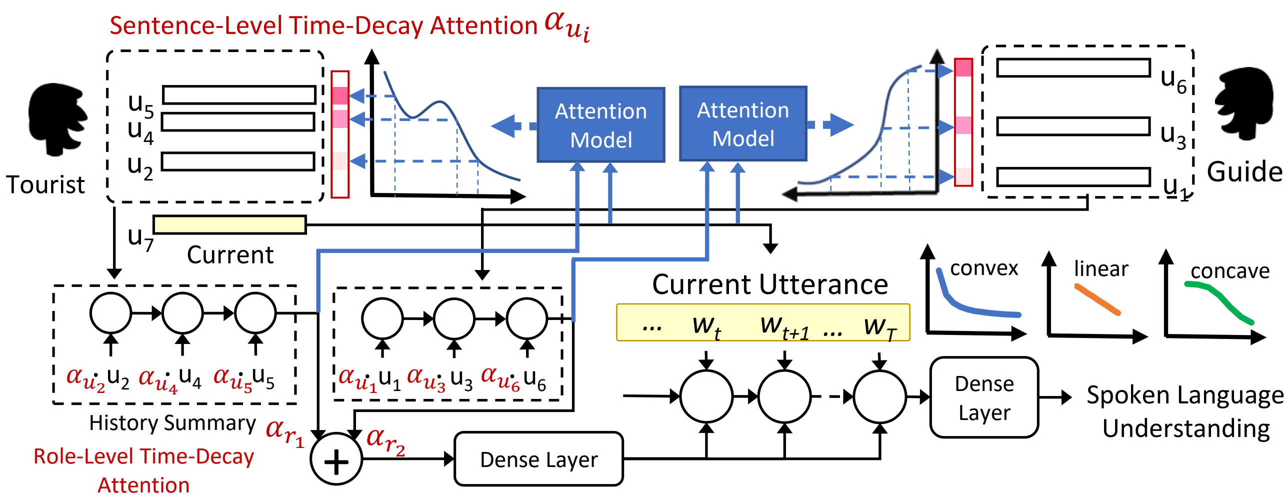


Role-Based Time-Decay Attention (Su+, 2018)

http://aclweb.org/anthology/N18-1194



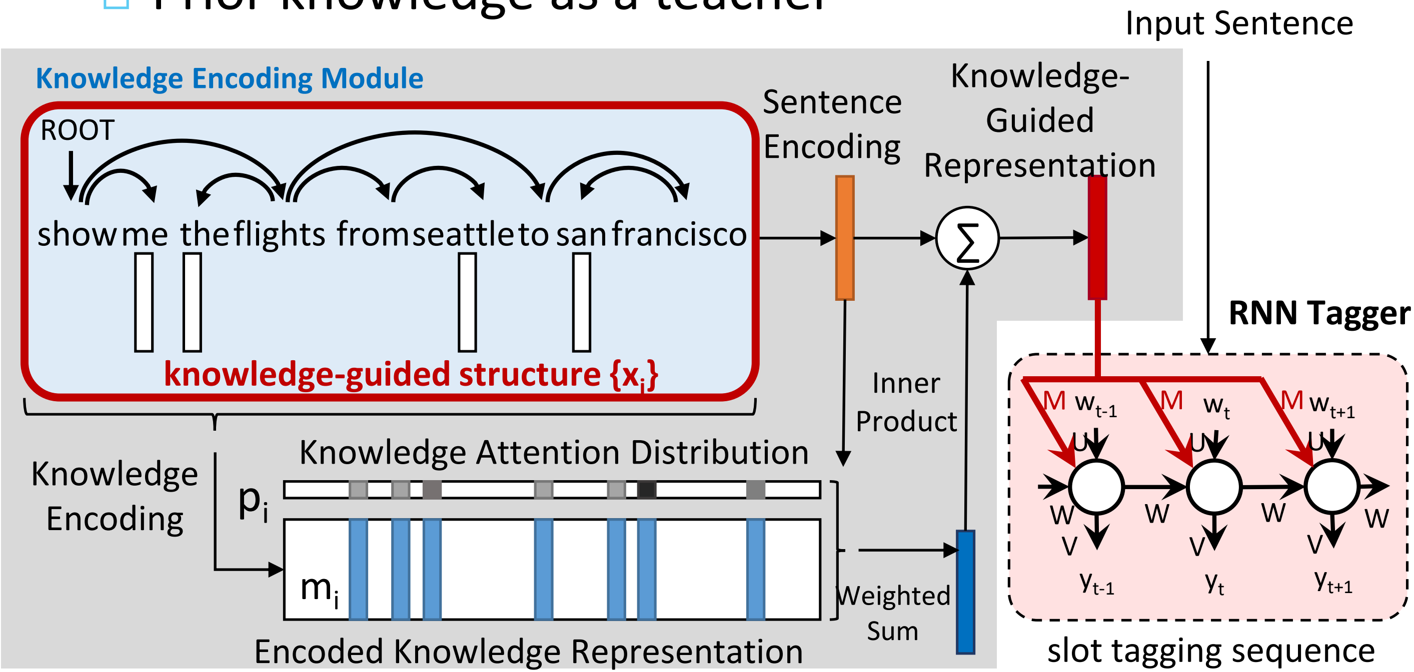
Context-Sensitive Time-Decay (Su+, 2018)



Time-decay attention significantly improves the understanding results

Structural LU (Chen+, 2016)

- Prior knowledge as a teacher



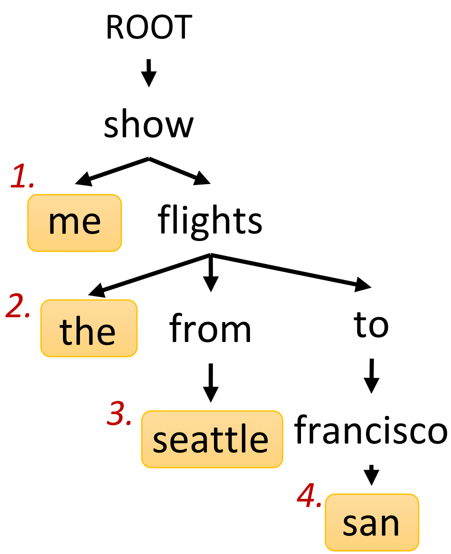
Structural LU (Chen+, 2016)

http://arxiv.org/abs/1609.03286

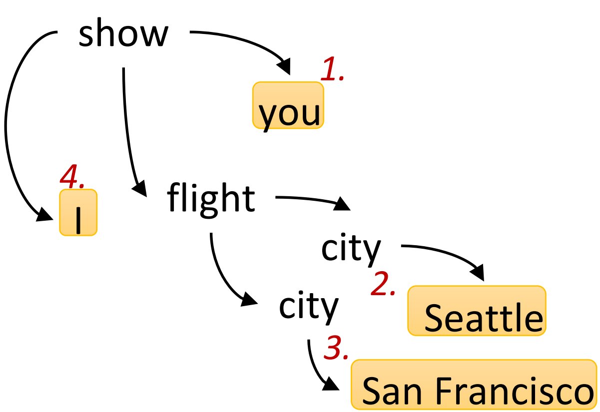
- Sentence structural knowledge stored as memory

Sentence s show me the flights from seattle to san francisco

Syntax (Dependency Tree)



Semantics (AMR Graph)



Structural LU (Chen+, 2016)

http://arxiv.org/abs/1609.03286

- Sentence structural knowledge stored as memory

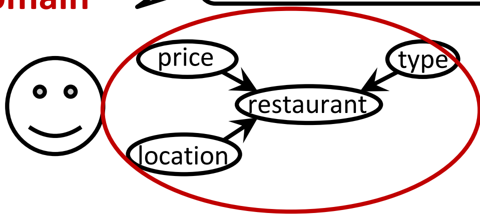
Using *less training data* with K-SAN allows the model pay the similar attention to the *salient substructures*

Semantic Frame Representation

- Requires a domain ontology: early connection to backend
- Contains **core content (intent, a set of slots with fillers)**

Restaurant Domain

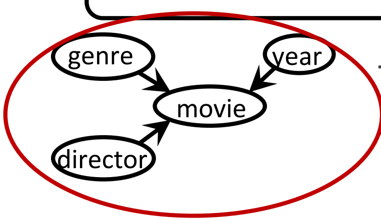
find me a cheap taiwanese restaurant in oakland



find_restaurant (price="cheap", type="taiwanese", location="oakland")

Movie Domain

show me action movies directed by james cameron



find_movie (genre="action", director="james cameron")

LU – Learning Semantic Ontology (Chen+, 2013)

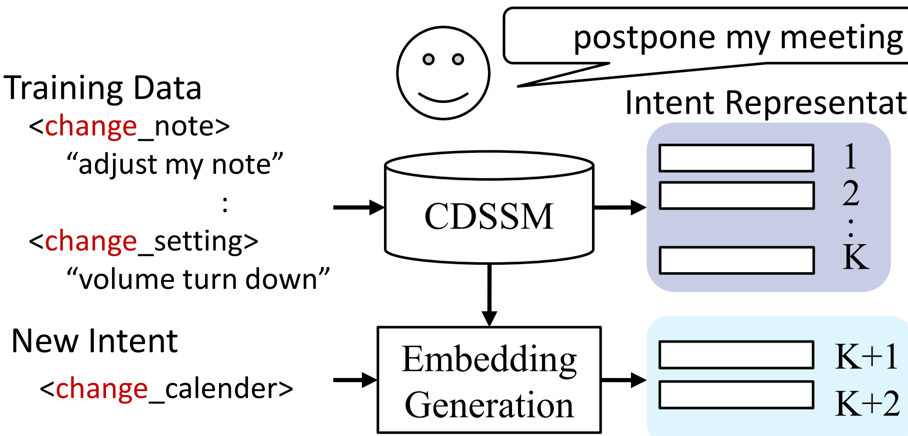
http://www.cs.cmu.edu/~ananlada/ConceptIdentificationICSLP02.pdf, http://ieeexplore.ieee.org/abstract/document/6707716/

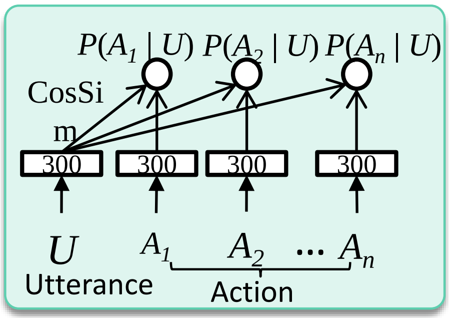
- Learning key domain concepts from goal-oriented human-human conversations
  - Clustering with mutual information and KL divergence (Chotimongkol & Rudnicky, 2002)
  - Spectral clustering based slot ranking model (Chen et al., 2013)
    - Use a state-of-the-art frame-semantic parser trained for FrameNet
    - Adapt the generic output of the parser to the target semantic space

LU – Intent Expansion (Chen+, 2016)

http://ieeexplore.ieee.org/abstract/document/7472838/

- Transfer dialogue acts across domains
  - Dialogue acts are similar for multiple domains
  - Learning new intents by information from other domains





The dialogue act representations can be automatically learned for other domains

LU – Language Extension (Upadhyay+, 2018)

http://shyamupa.com/papers/UFTHH18.pdf

- 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

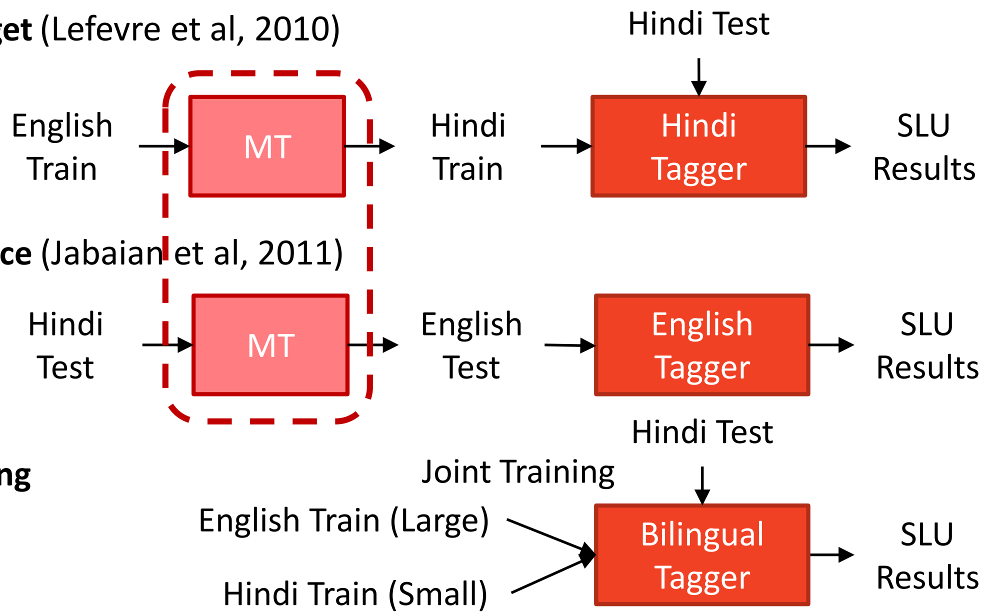
LU – Language Extension (Upadhyay+, 2018)

http://shyamupa.com/papers/UFTHH18.pdf

Train on Target (Lefevre et al, 2010)

Test on Source (Jabaian et al, 2011)

Joint Training

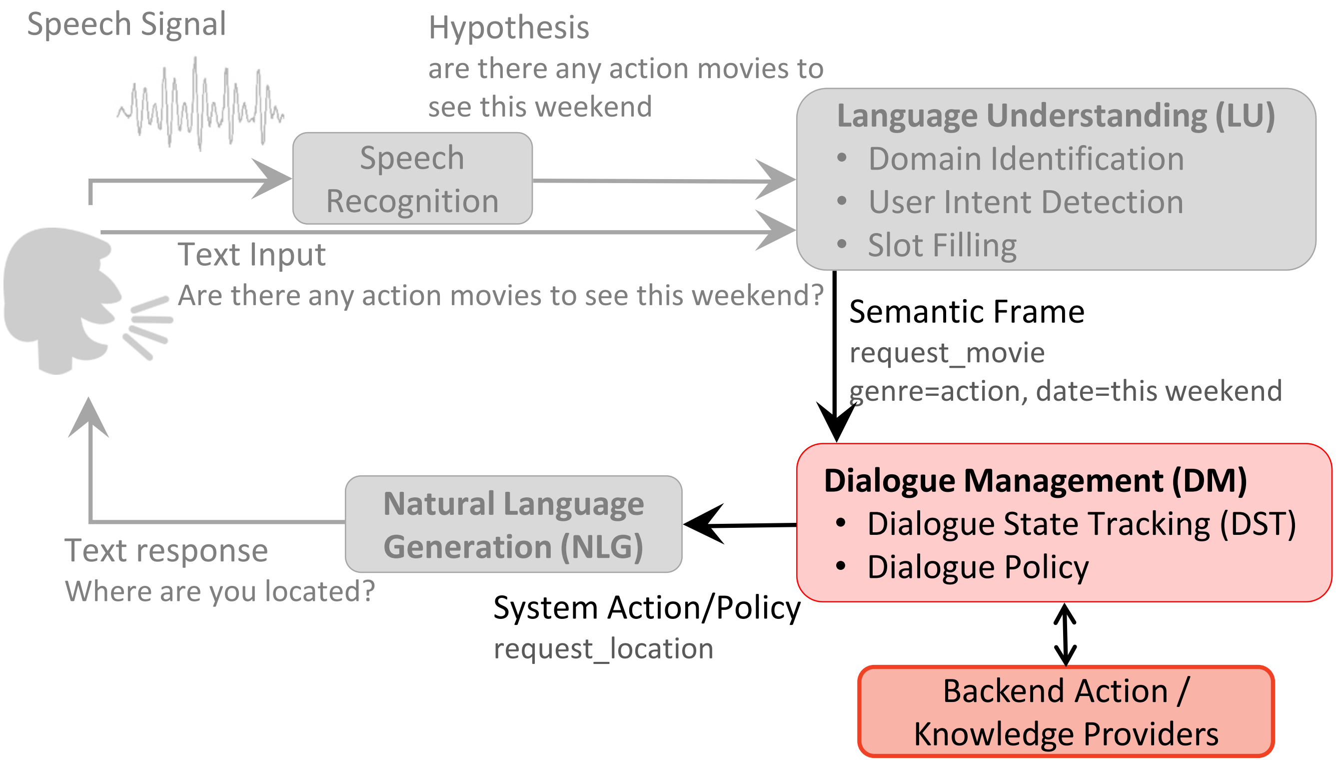


MT system is not required and both languages can be processed by a single model

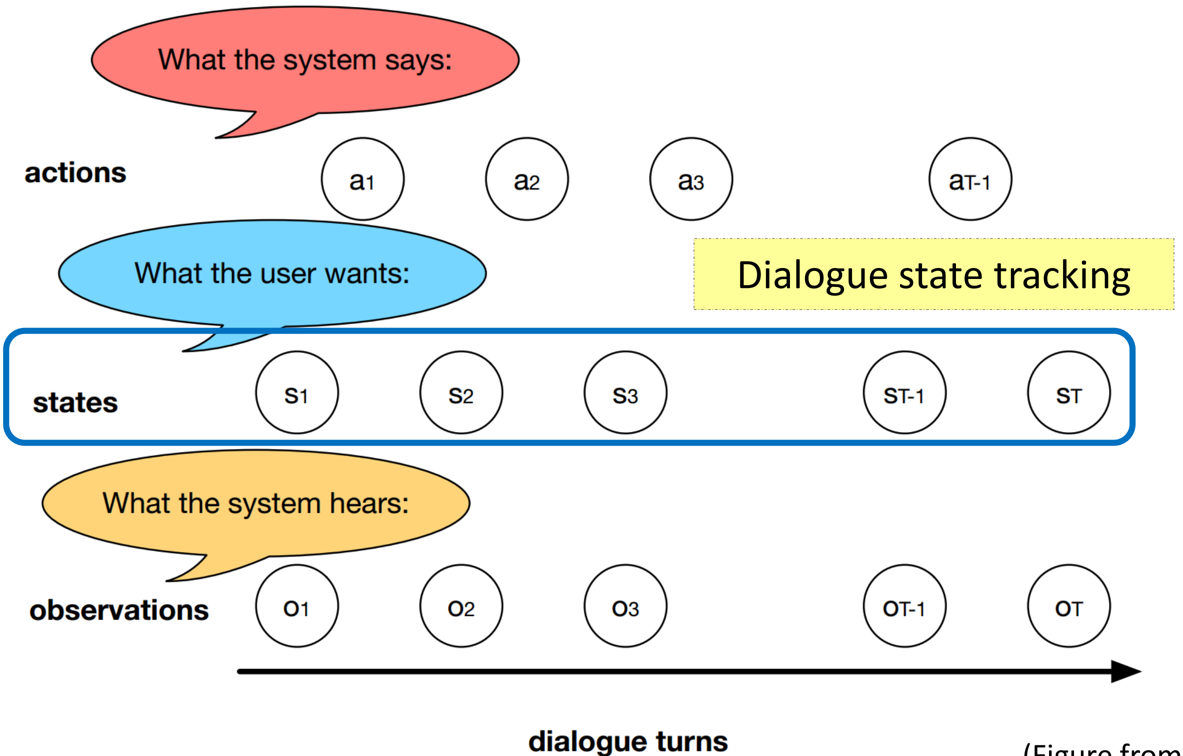
Outline

- Introduction & Background
  - Neural Networks
  - Reinforcement Learning
- Modular Dialogue System
  - Spoken/Natural Language Understanding (SLU/NLU)
  - **Dialogue Management (DM)**
    - **Dialogue State Tracking (DST)**
    - Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - End-to-End Neural Dialogue Systems
- System Evaluation
- Recent Trends on Learning Dialogues

Task-Oriented Dialogue System (Young, 2000)



Elements of Dialogue Management



(Figure from Gašić)

Dialogue State Tracking (DST)

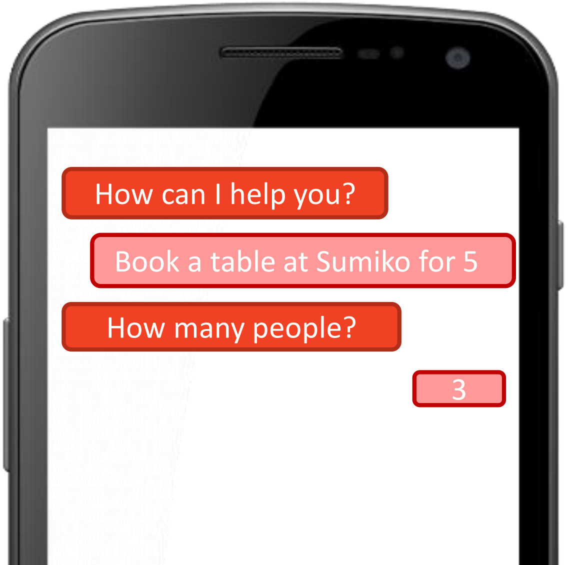
- Dialogue state: *a representation of the system's belief of the user's goal(s) at any time during the dialogue*
- Inputs
  - Current user utterance
  - Preceding system response
  - Results from previous turns
- For
  - Looking up knowledge or making API call(s)
  - Generating the next system action/response

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)



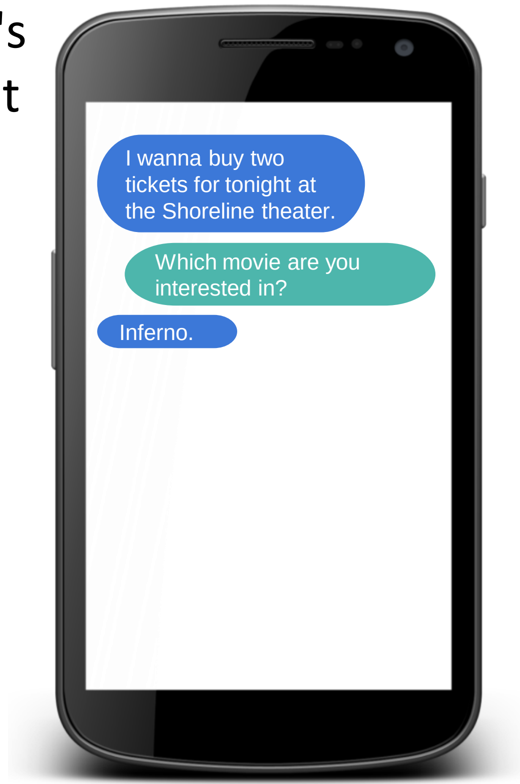
Multi-Domain Dialogue State Tracking

- 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

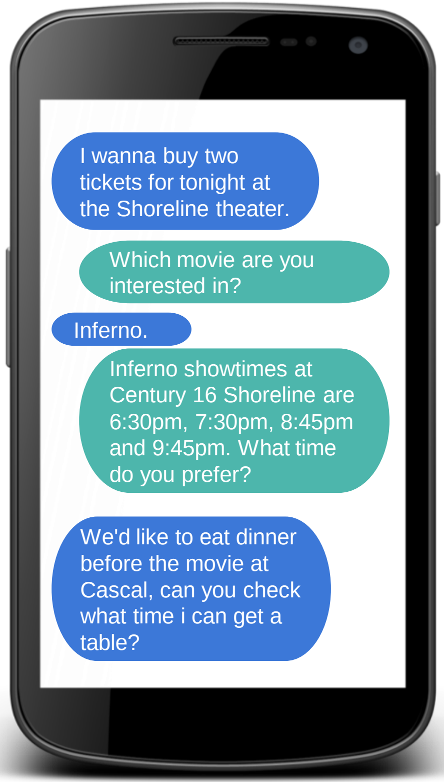
Movies

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

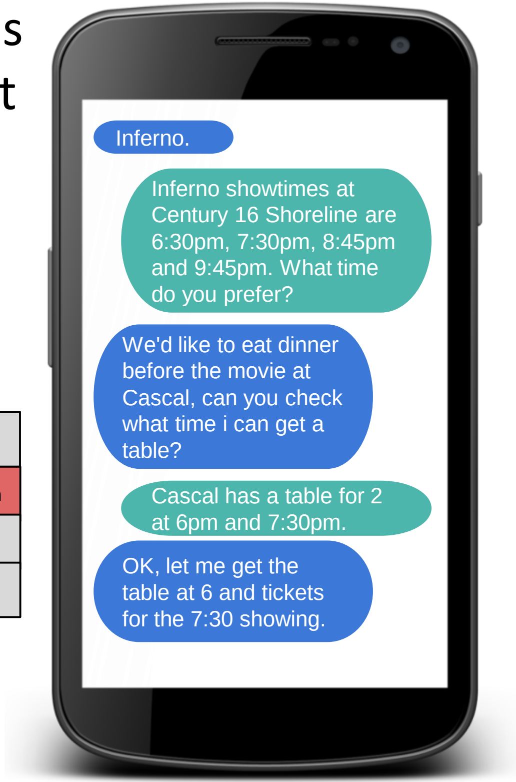
Movies

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

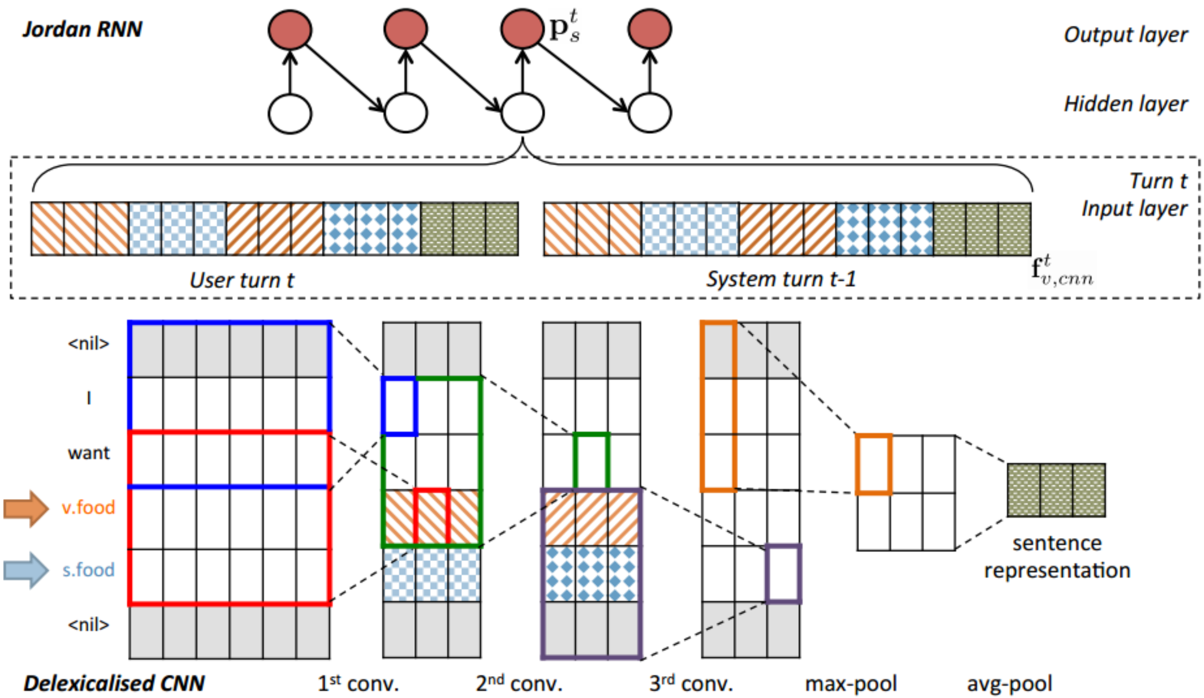
Date	11/15/17		
Time	6:00 pm	6:30 pm	7:00 pm
Restaurant	Cascal		
#People	2		

Less Likely — More Likely



RNN-CNN DST (Mrkšić+, 2015)

https://arxiv.org/abs/1506.07190

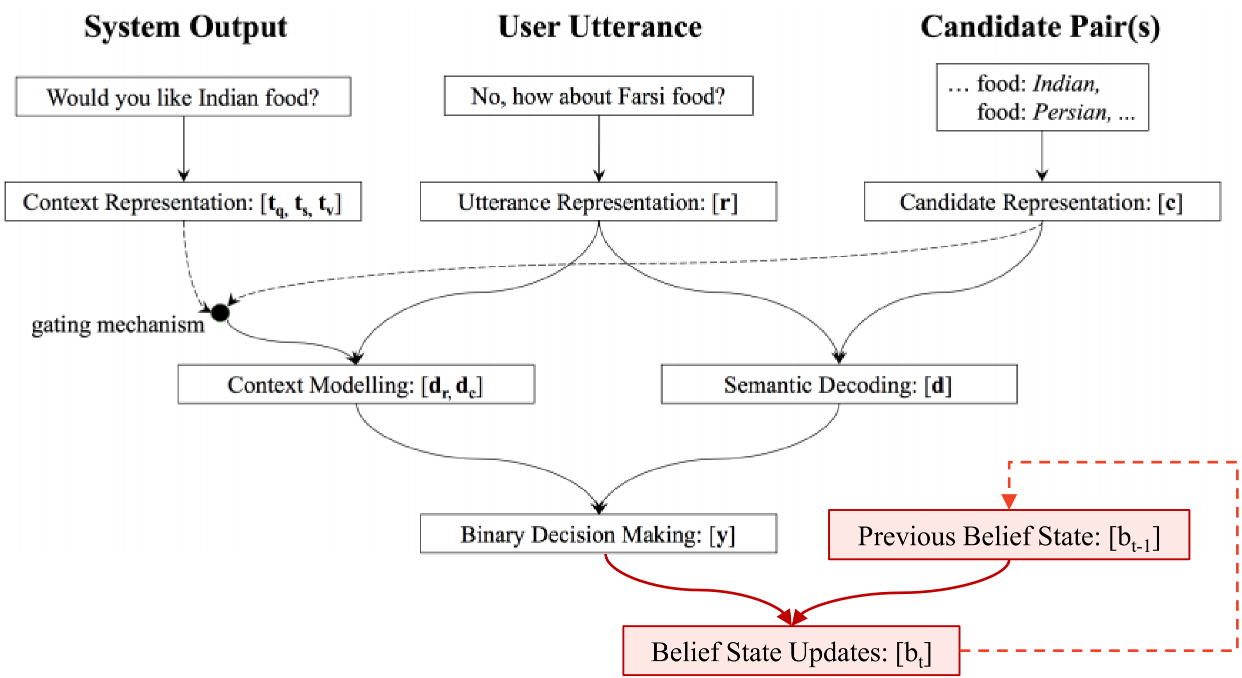


(Figure from Wen et al, 2016)

Neural Belief Tracker (Mrkšić+, 2016)

https://arxiv.org/abs/1606.03777

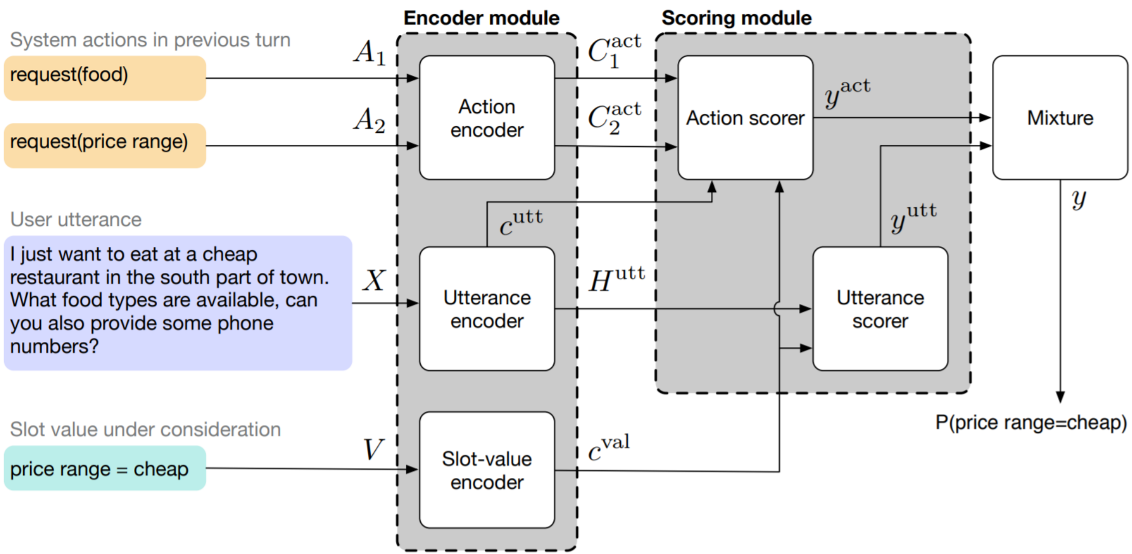
- Candidate pairs are considered



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

http://www.aclweb.org/anthology/P18-1135

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



Dialog State Tracking Challenge (DSTC)

(Williams+, 2013, Henderson+, 2014, Henderson+, 2014, Kim+, 2016, Kim+, 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

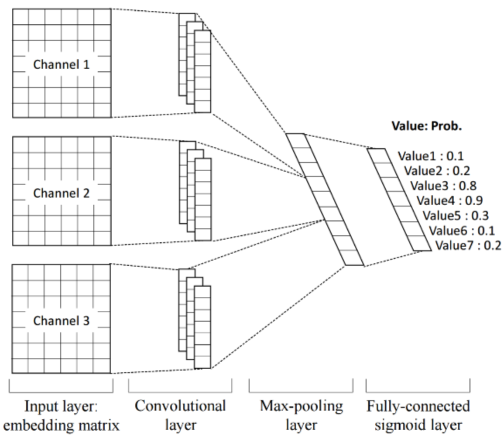
DST Evaluation

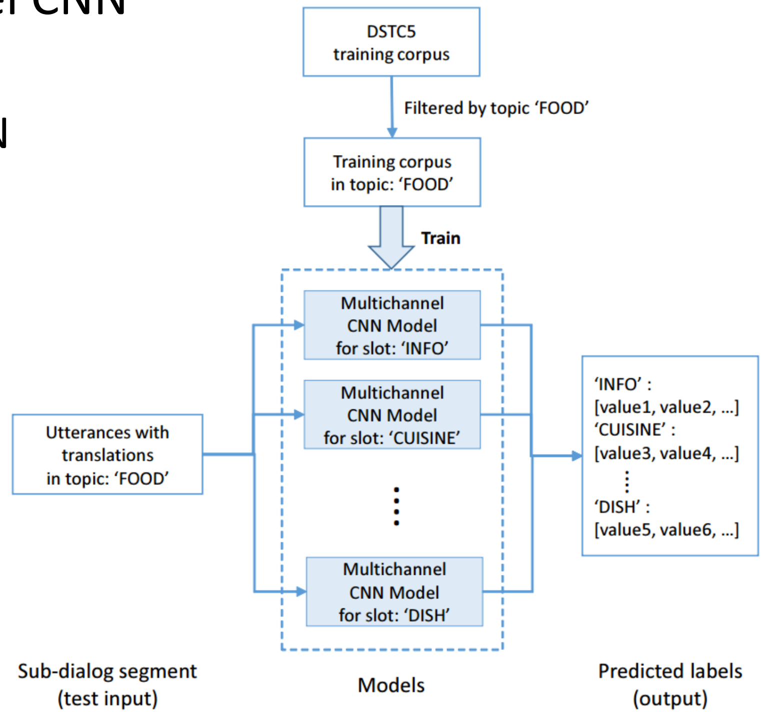
- Metric
  - Tracked state accuracy with respect to user goal
  - Recall/Precision/F-measure individual slots

DST – Language Extension (Shi+, 2016)

https://arxiv.org/abs/1701.06247

- Training a multichannel CNN for each slot
  - Chinese character CNN
  - Chinese word CNN
  - English word CNN





DST – Task Lineages (Lee & Stent, 2016)

https://www.aclweb.org/anthology/W/W16/W16-36.pdf#page=29

- Slot values shared across tasks
- Utterances with complex constraints on user goals
- Interleaved multiple task discussions

Task Frame:

Connection to Manhattan and find me a Thai restaurant, not Italian

$$\begin{bmatrix} \begin{bmatrix} \textbf{Task} & \text{Transit} \\ \textbf{DAIs} & (0.8, \text{inform}(\text{dest}=\text{MH})_{0.7}^{0.1}) \end{bmatrix} \\ \begin{bmatrix} \textbf{Task} & \text{Restaurant} \\ \textbf{DAIs} & (0.7, \text{inform}(\text{food}=\text{thai})_{1.2}^{0.9}) \\ & (0.6, \text{deny}(\text{food}=\text{italian})_{1.7}^{1.4}) \end{bmatrix} \end{bmatrix}$$

(confidence, dialog act item$^{\text{Start_time}}_{\text{End_time}}$)

Task State:

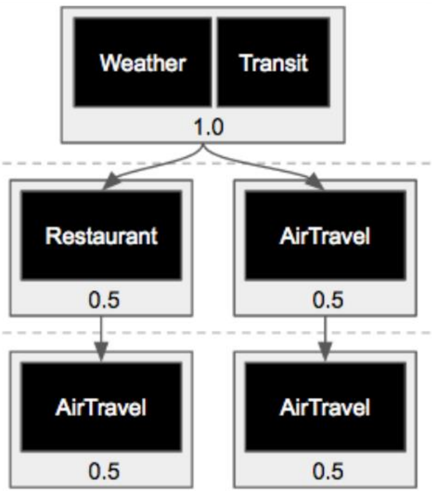
Thai restaurant, not Italian

$$\begin{bmatrix} \textbf{Task} & \text{Restaurant} \\ \textbf{Constraints} & (0.7, \text{food} = \text{thai}) \\ & (0.6, \text{food} \neq \text{italian}) \\ \textbf{DB} & [\text{"Thai To Go", "Pa de Thai"}] \\ \textbf{Timestamps} & 01/01/2016 : 12\text{-}00\text{-}00 \\ \dots & \dots \end{bmatrix}$$

DST – Task Lineages (Lee & Stent, 2016)

https://www.aclweb.org/anthology/W/W16/W16-36.pdf#page=29

Turn	User Input	Task Lineages
0	Weather in New York. Connection to Manhattan	
1	Want to go to Thai	
2	I want to travel to Thai	



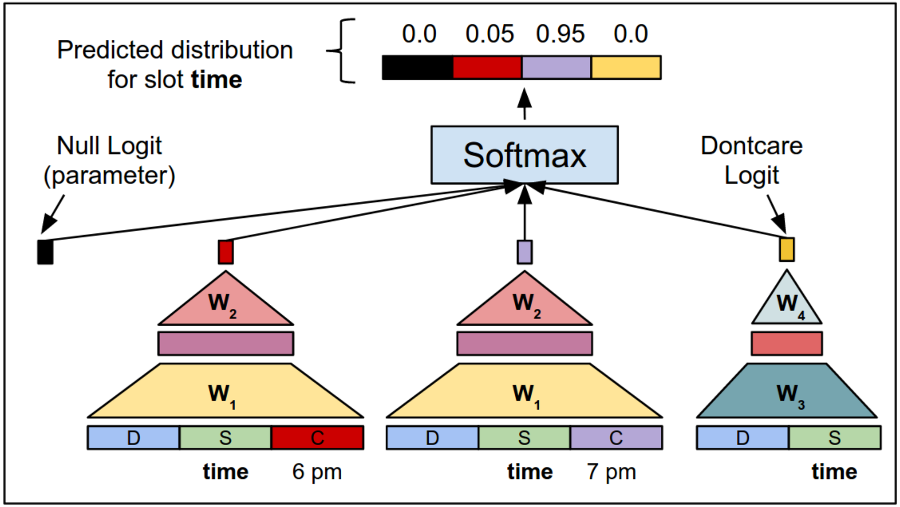
DST – Scalability (Rastogi+, 2017)

https://arxiv.org/pdf/1712.10224.pdf

- Focus only on the relevant slots
- Better generalization to ASR lattices, visual context, etc.

S> How about 6 pm?

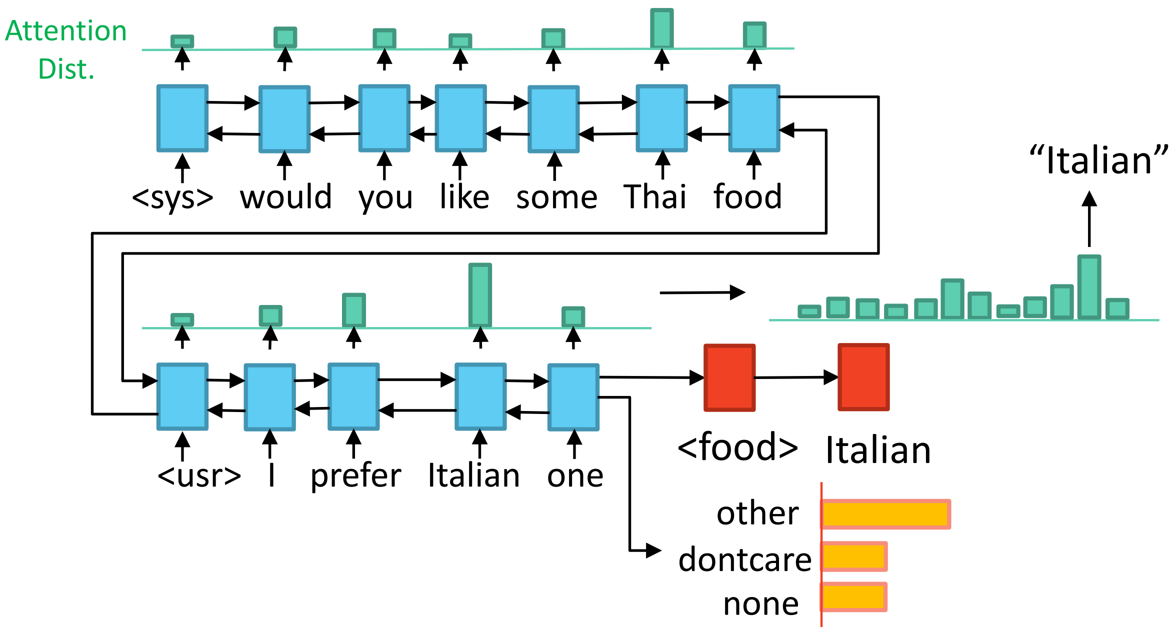
U> I am busy then, book it for 7 pm instead.



DST – Handling Unknown Values (Xu & Hu, 2018)

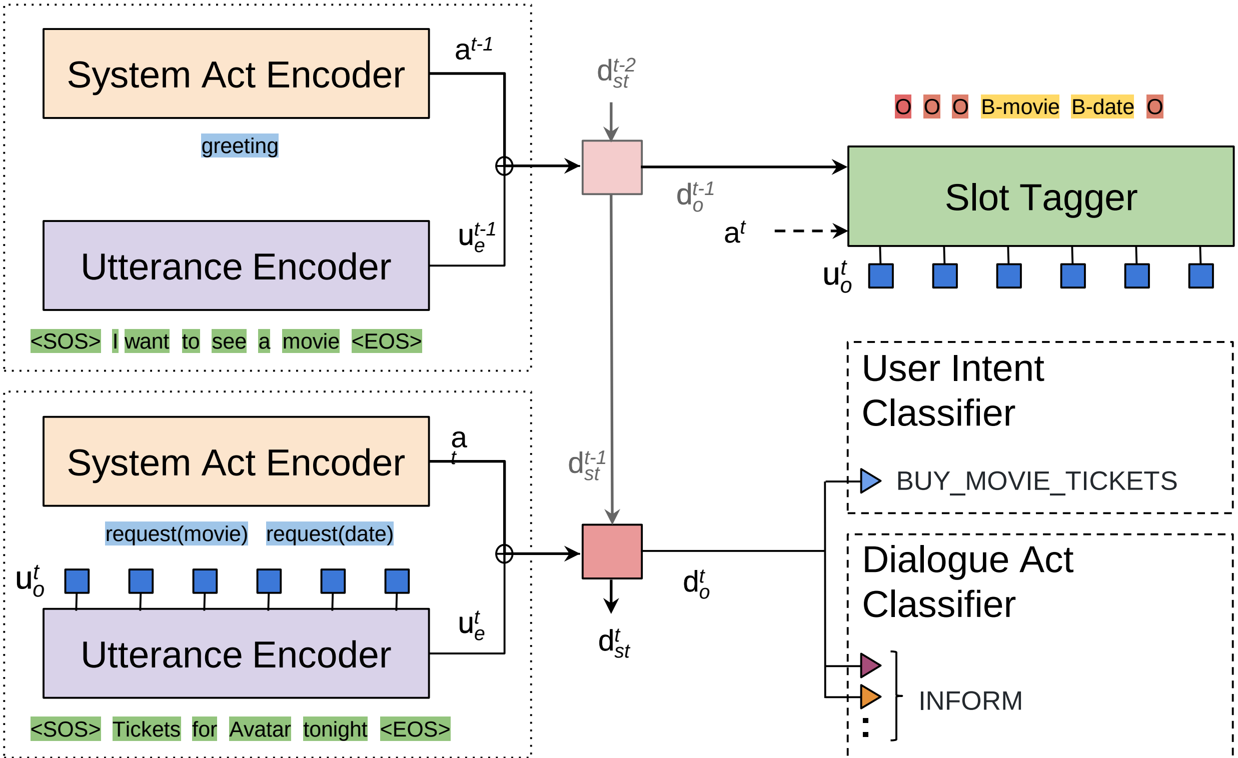
http://aclweb.org/anthology/P18-1134

- Issue: fixed value sets in DST

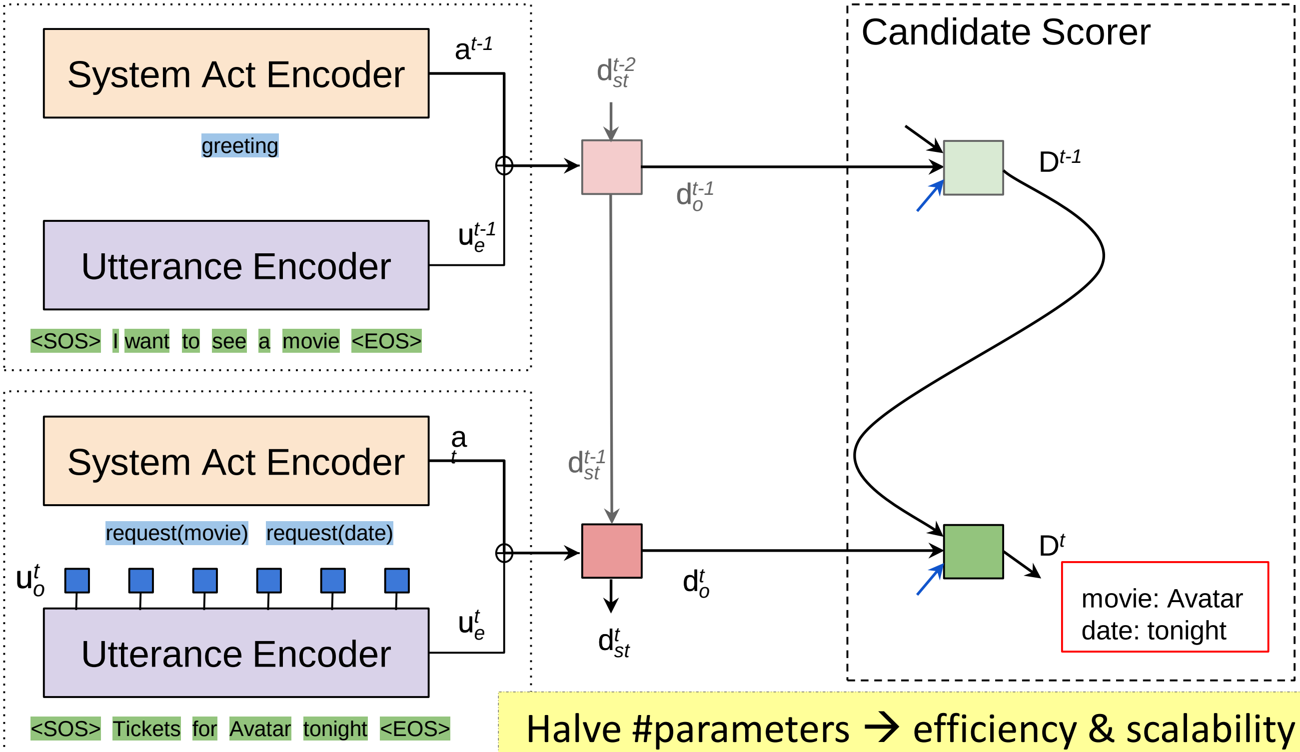


Pointer networks for generating unknown values

Joint NLU and DST (Gupta+, 2018)



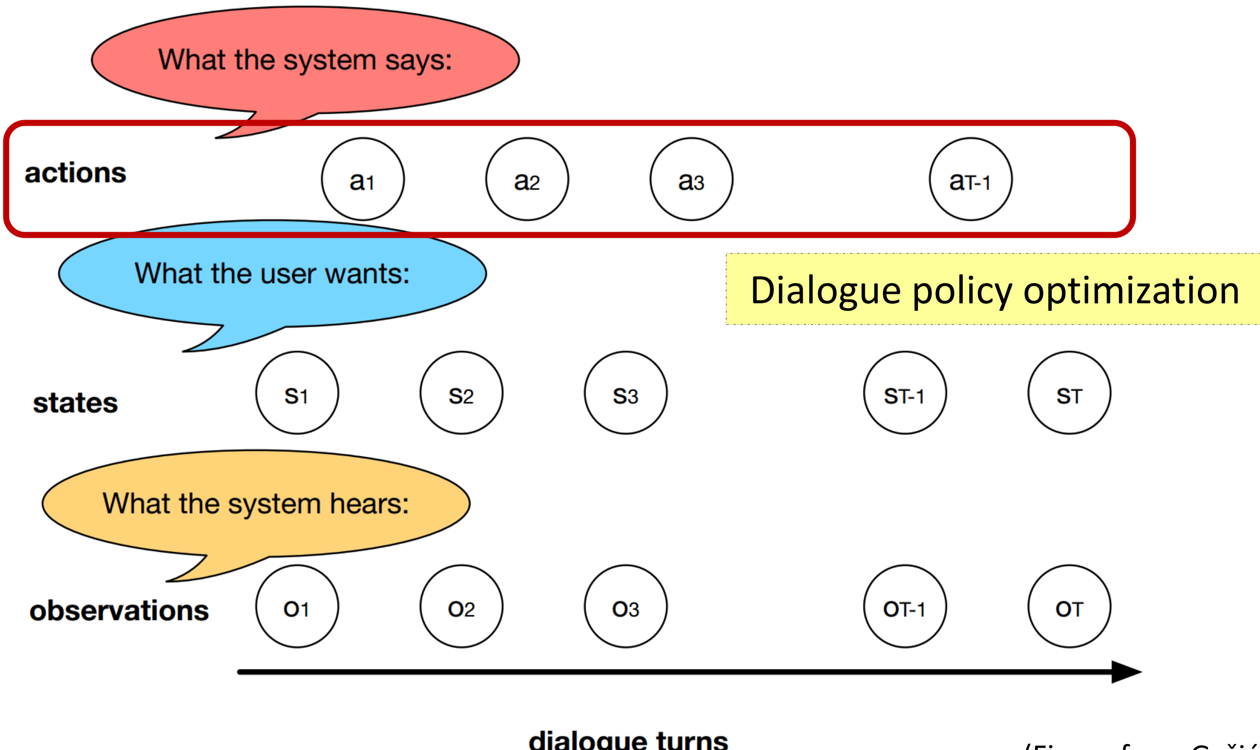
Joint NLU and DST (Gupta+, 2018)



Halve #parameters → efficiency & scalability

Outline

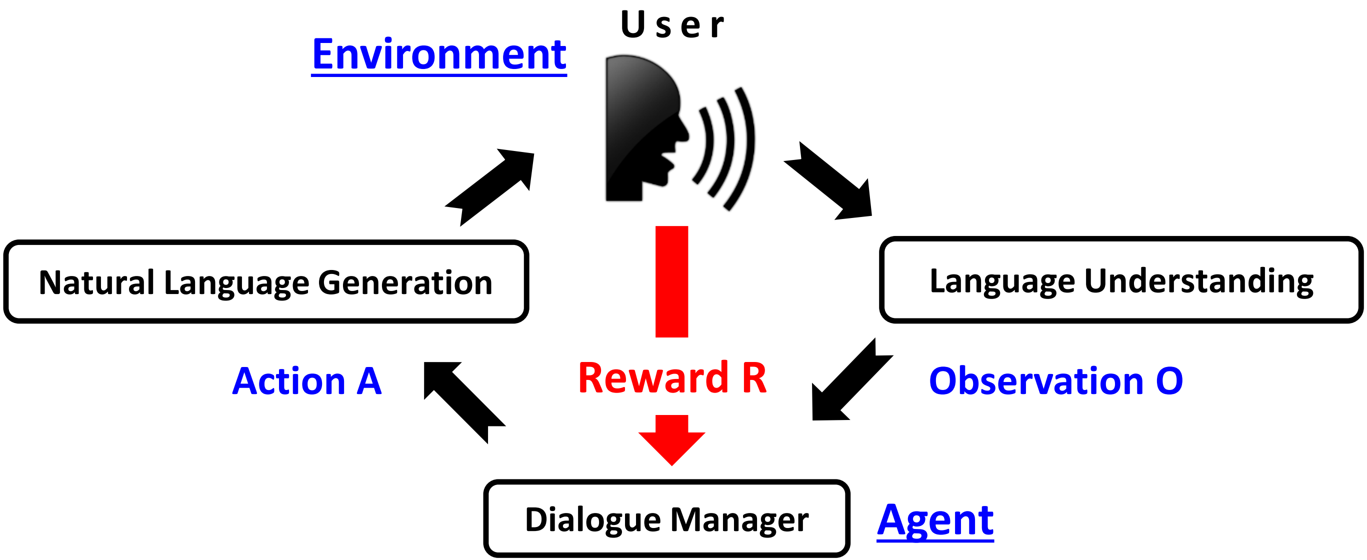
Elements of Dialogue Management



(Figure from Gašić)

Dialogue Policy Optimization

- Dialogue management in a RL framework



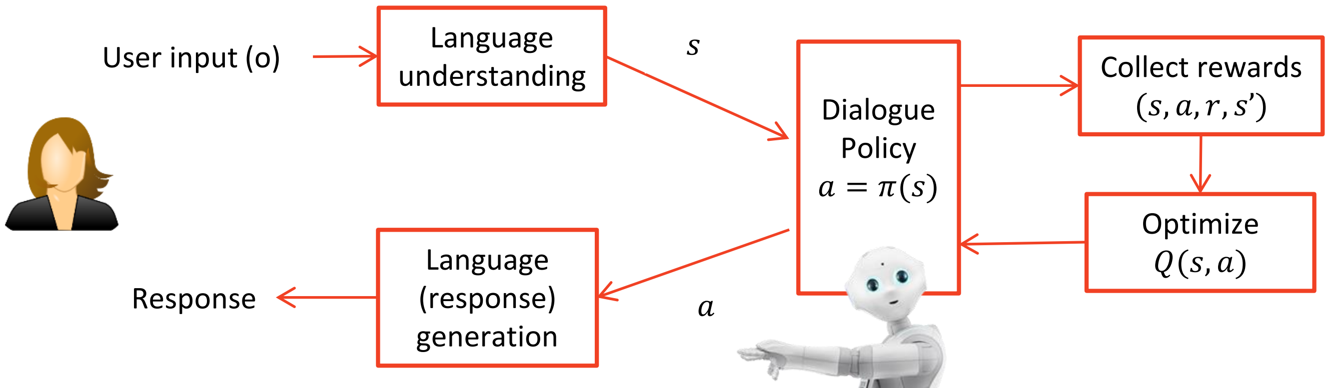
Goal: select the best action that maximizes the future reward

Reward for RL $\cong$ 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

RL for Dialogue Policy Optimization



Type of Bots	State	Action	Reward
Social ChatBots	Chat history	System Response	# of turns maximized; Intrinsically motivated reward
InfoBots (interactive Q/A)	User current question + Context	Answers to current question	**Relevance of answer; # of turns minimized**
Task-Completion Bots	User current input + Context	System dialogue act w/ slot value (or API calls)	**Task success rate; # of turns minimized**

Goal: develop a generic deep RL algorithm to learn dialogue policy for all bot categories

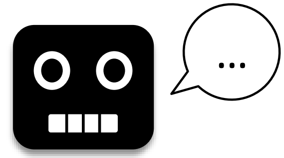
Dialogue Reinforcement Learning Signal

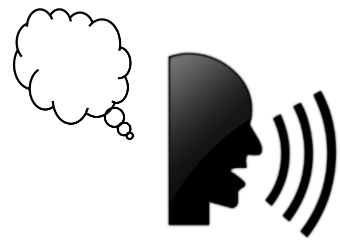
Typical reward function

- Large reward at completion if successful
- -1 for per turn penalty

Typically requires domain knowledge

- ✔ Simulated user
- ✔ Paid users (Amazon Mechanical Turk)
- ✘ Real users



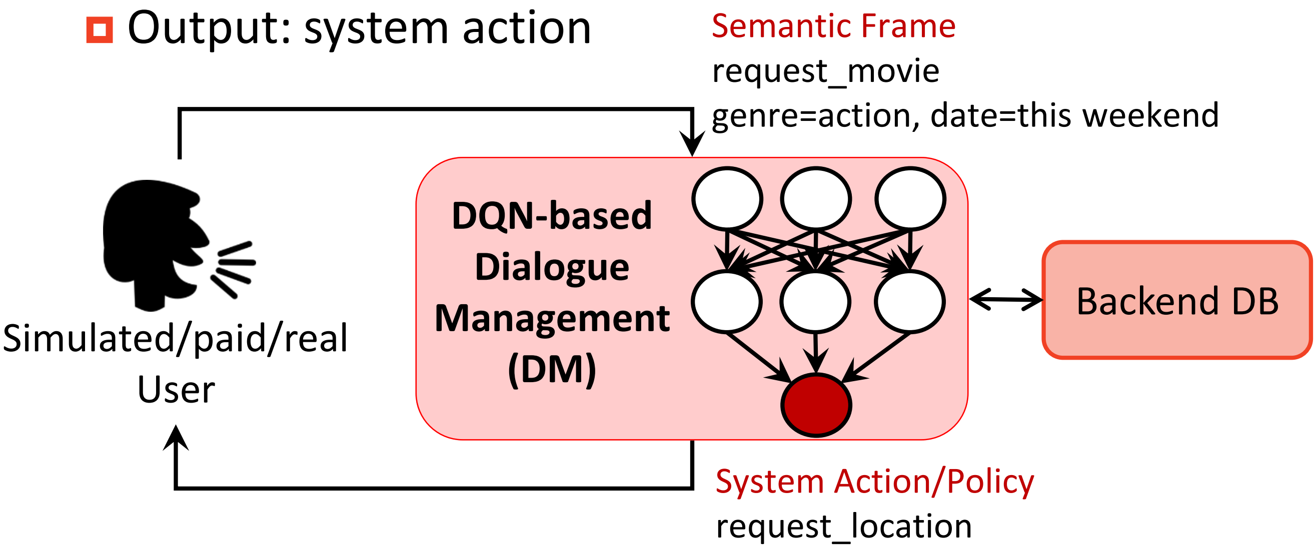


The **user simulator** is usually required for dialogue system training before deployment

Neural Dialogue Manager (Li+, 2017)

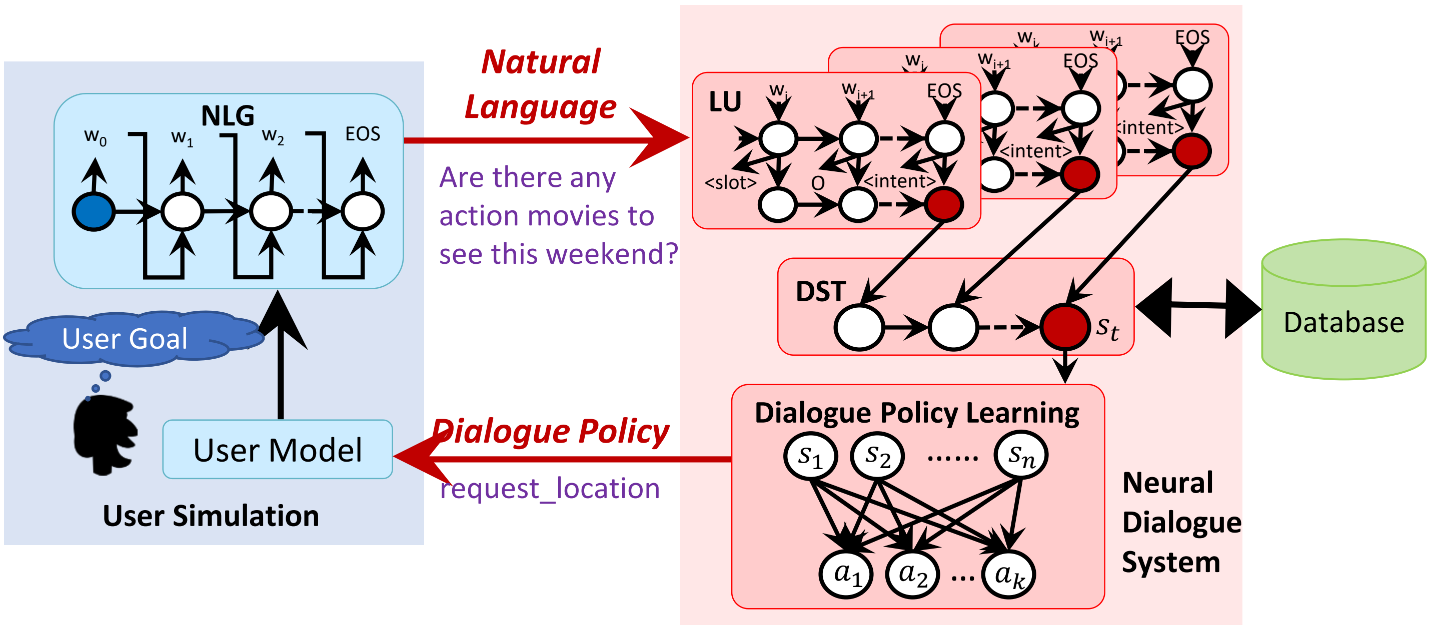
http://www.aclweb.org/anthology/I17-1074

- Deep RL for training DM
  - Input: current semantic frame observation, database returned results
  - Output: system action



E2E Task-Completion Bot (TC-Bot) (Li+, 2017)

http://www.aclweb.org/anthology/I17-1074

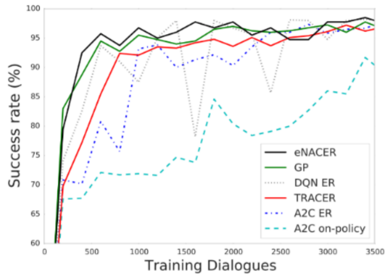


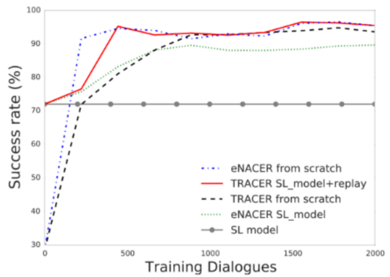
Idea: SL for each component and RL for end-to-end training

SL + RL for Sample Efficiency (Su+, 2017)

http://aclweb.org/anthology/W17-5518

- Issue about RL for DM
  - slow learning speed
  - cold start
- Solutions
  - Sample-efficient actor-critic
    - Off-policy learning with experience replay
    - Better gradient update
  - Utilizing supervised data
    - Pretrain the model with SL and then fine-tune with RL
    - Mix SL and RL data during RL learning
    - Combine both

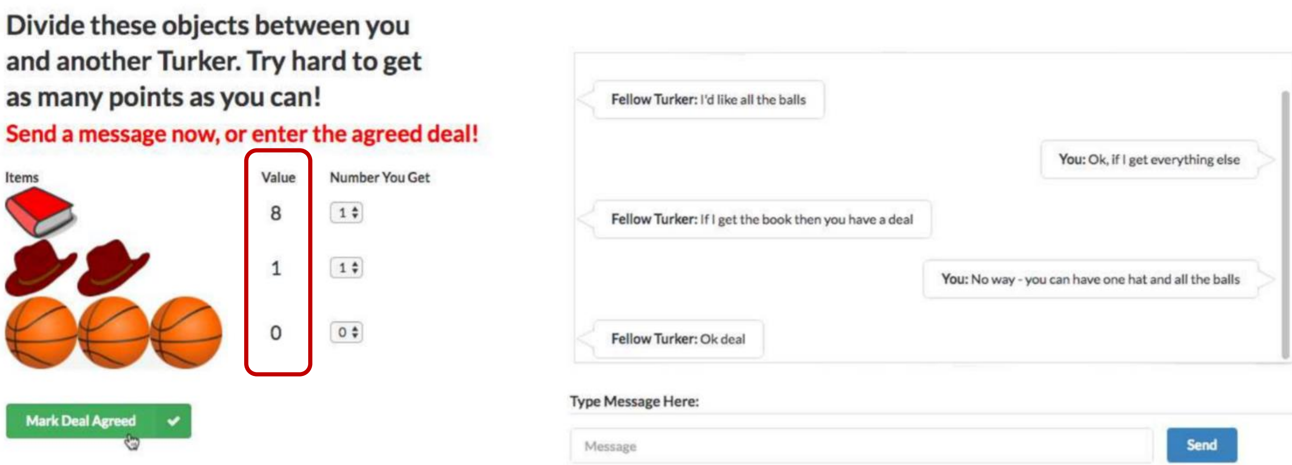




Learning to Negotiate (Lewis+, 2017)

https://arxiv.org/pdf/1706.05125.pdf

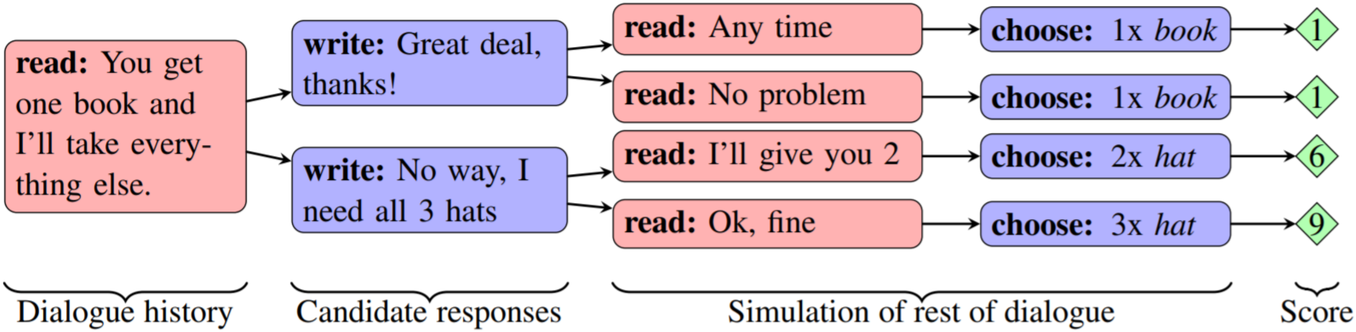
- Task: multi-issue bargaining
  - Each agent has its own value function



Learning to Negotiate (Lewis+, 2017)

https://arxiv.org/pdf/1706.05125.pdf

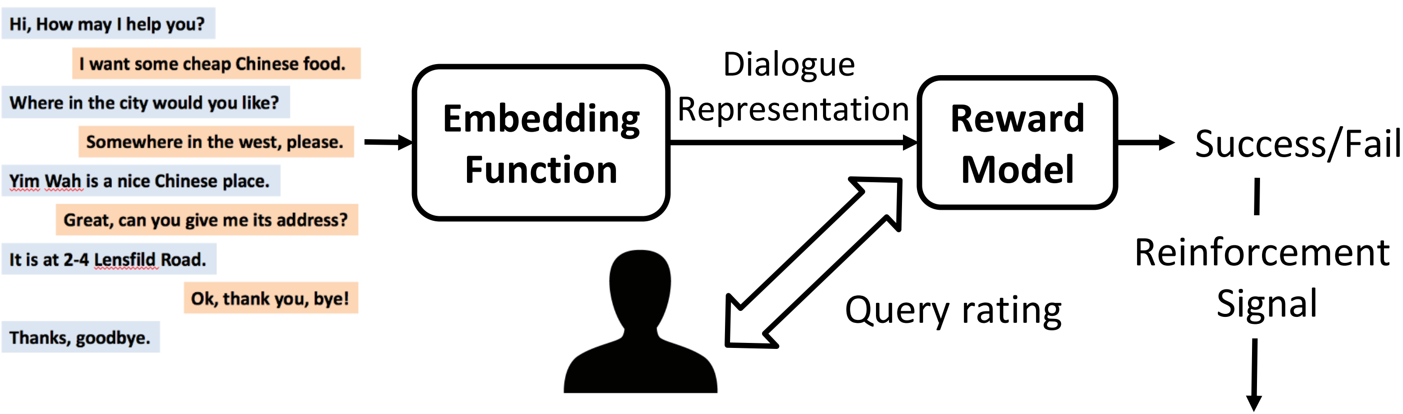
- Dialogue rollouts to simulate a future conversation
- SL + RL
  - SL aims to imitate human users' actions
  - RL tries to make agents focus on the goal



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

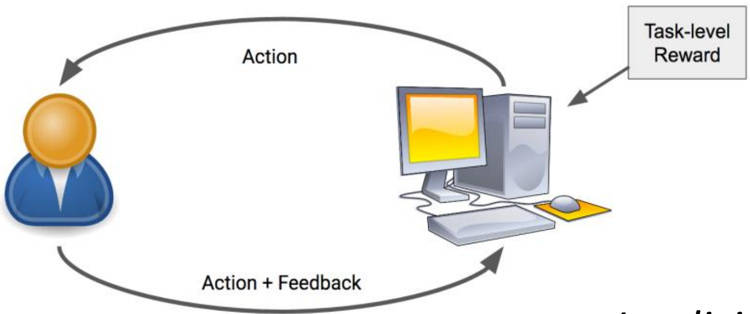
http://www.anthology.aclweb.org/W/W15/W15-46.pdf#page=437; https://www.aclweb.org/anthology/P/P16/P16-1230.pdf

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



Interactive RL for DM (Shah+, 2016)

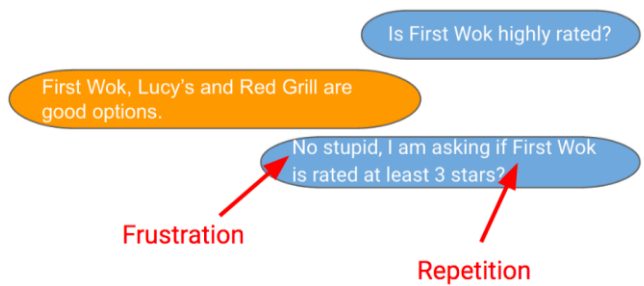
https://research.google.com/pubs/pub45734.html



Immediate Feedback

Explicit

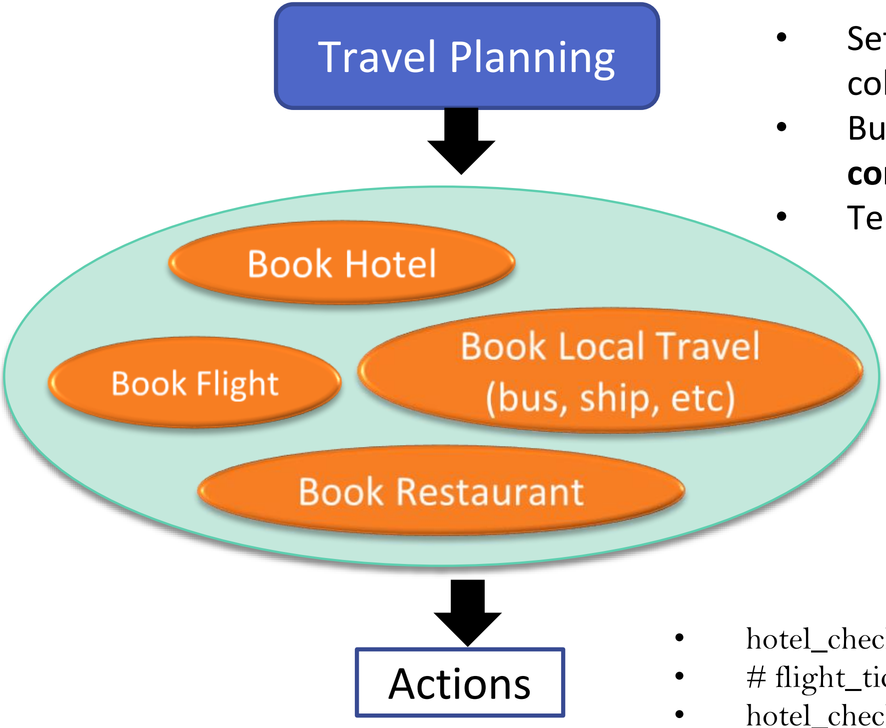
Implicit



Use a third agent for providing interactive feedback to the DM

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

https://arxiv.org/abs/1704.03084



- 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,

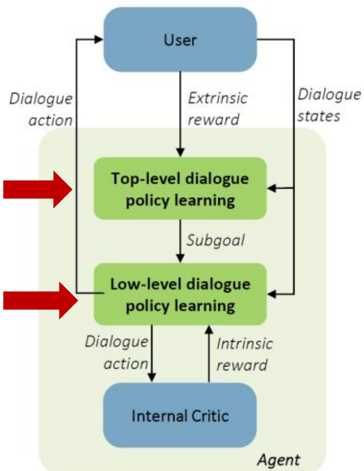
Multi-Domain – Hierarchical RL (Peng+, 2017)

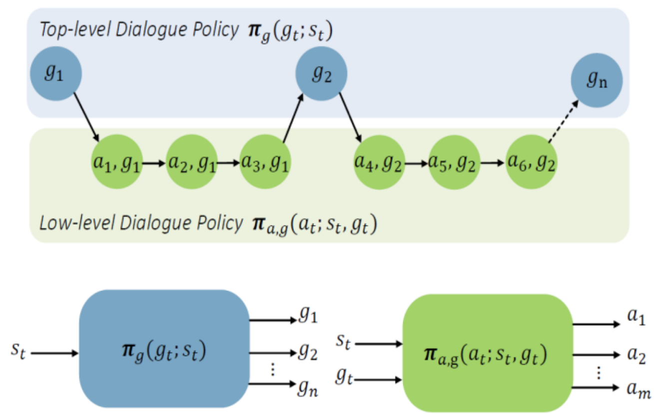
https://arxiv.org/abs/1704.03084

- Model makes decisions over two levels: *meta-controller & controller*
- The *agent* learns these policies simultaneously
  - the 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



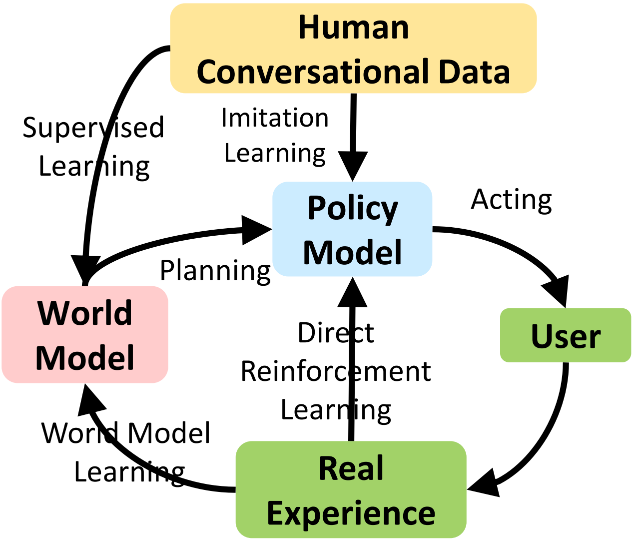


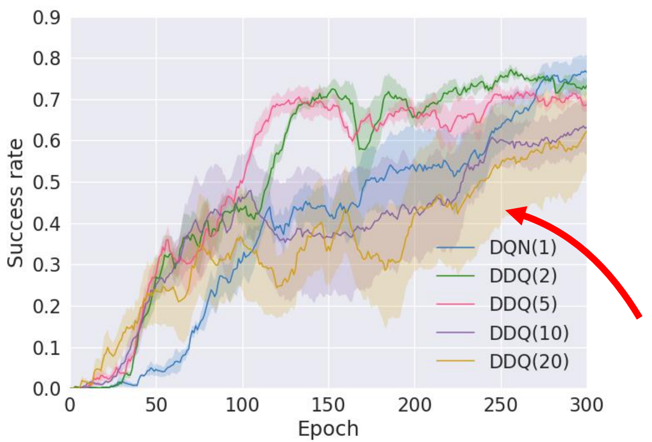
(mitigate reward sparsity issues)

Planning – Deep Dyna-Q (Peng+, 2018)

https://arxiv.org/abs/1801.06176

- Issues: sample-inefficient, discrepancy between simulator & real user
- Idea: learning with real users with planning



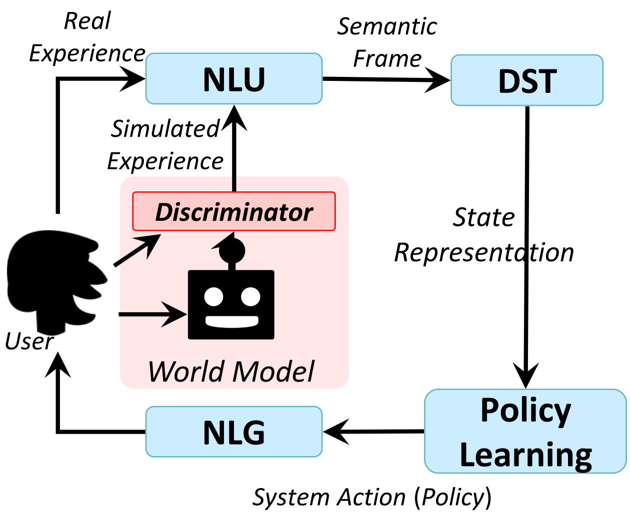


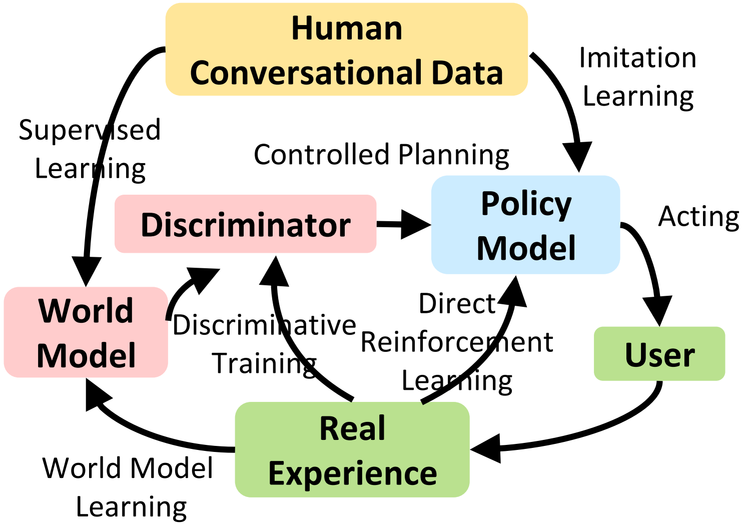
Policy learning suffers from the poor quality of fake experiences

Robust Planning – D3Q: Discriminative Deep Dyna-Q (Su+, 2018)

(to appear) EMNLP 2018

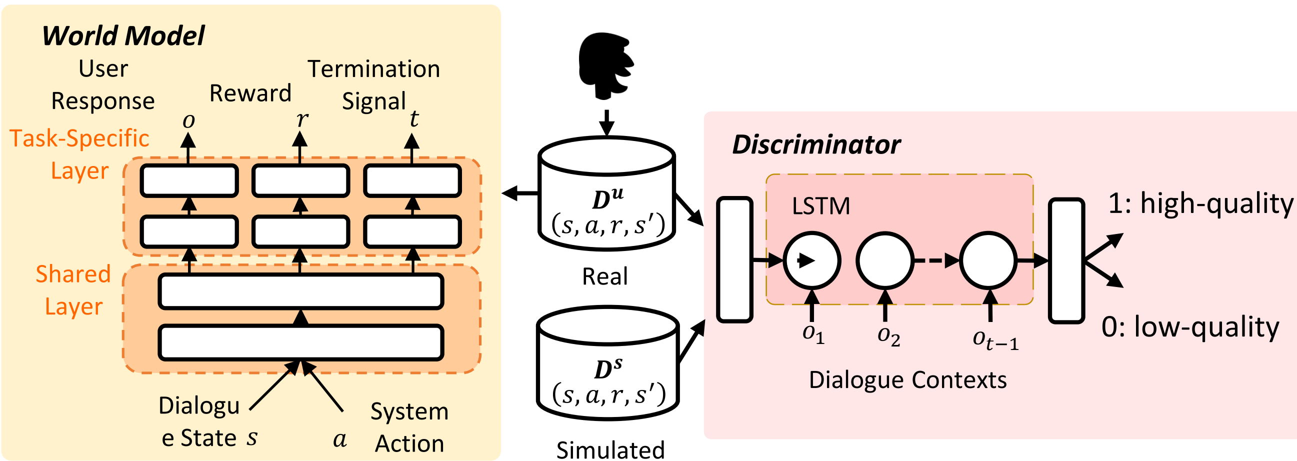
- Idea: add a *discriminator* to filter out the bad experiences





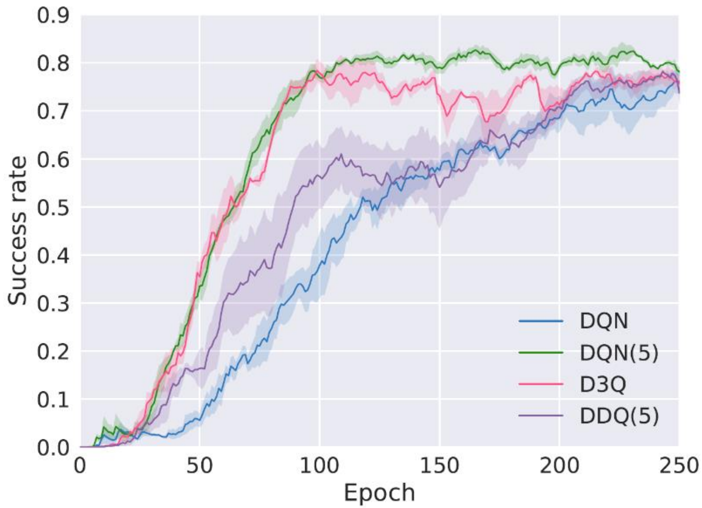
Robust Planning – D3Q: Discriminative Deep Dyna-Q (Su+, 2018)

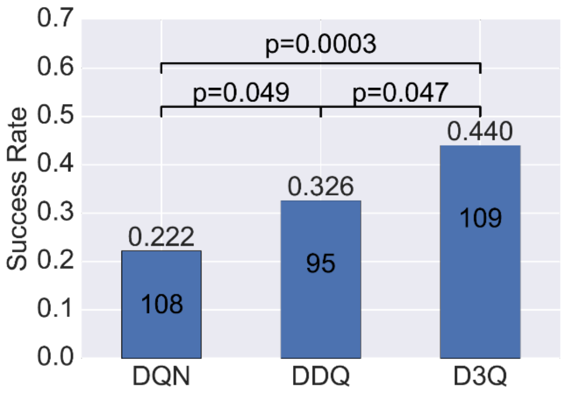
(to appear) EMNLP 2018



Robust Planning – D3Q: Discriminative Deep Dyna-Q (Su+, 2018)

(to appear) EMNLP 2018





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

Dialogue Management Evaluation

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

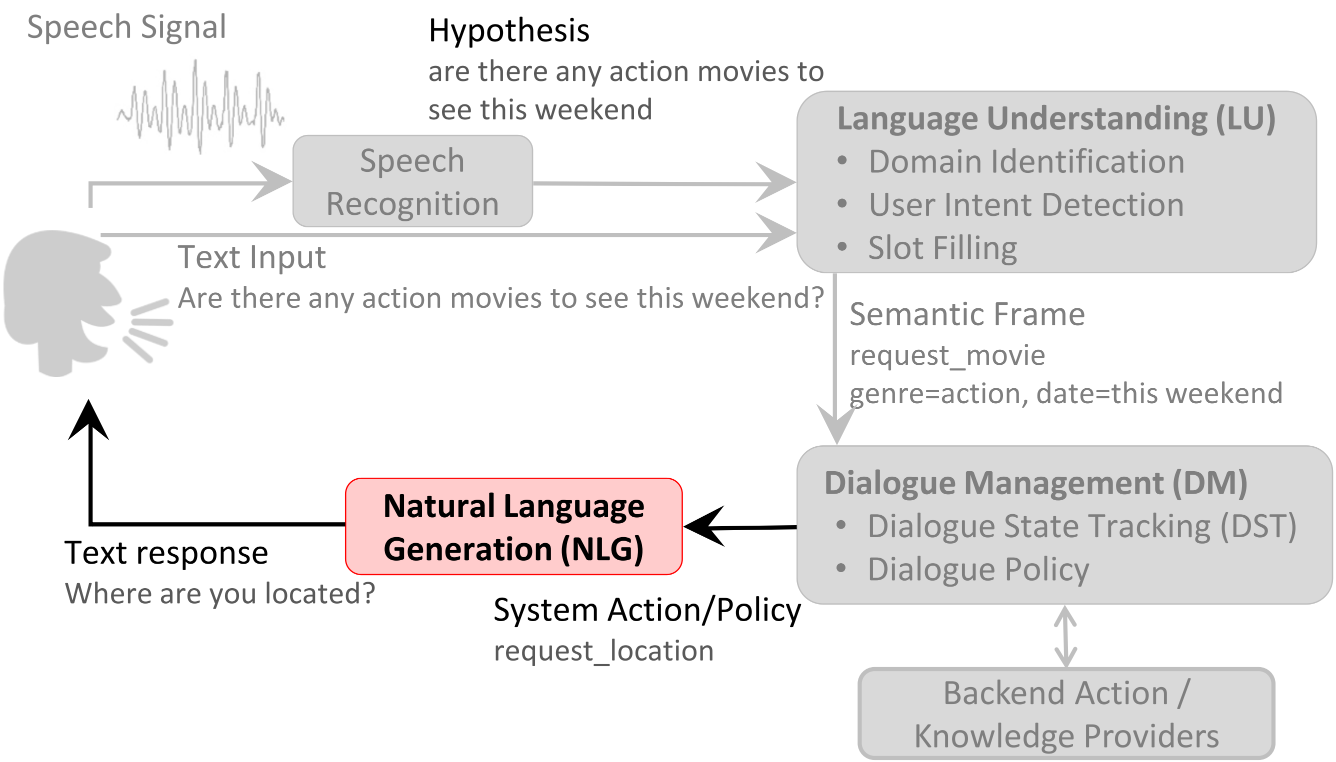
RL-Based DM Challenge

- SLT 2018 Microsoft Dialogue Challenge: End-to-End Task-Completion Dialogue Systems
  - Domain 1: Movie-ticket booking
  - Domain 2: Restaurant reservation
  - Domain 3: Taxi ordering

Outline

- Introduction & Background
  - Neural Networks
  - Reinforcement Learning
- Modular Dialogue System
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management (DM)
    - Dialogue State Tracking (DST)
    - Dialogue Policy Optimization
  - **Natural Language Generation (NLG)**
  - End-to-End Neural Dialogue Systems
- System Evaluation
- Recent Trends on Learning Dialogues

Task-Oriented Dialogue System (Young, 2000)



Natural Language Generation (NLG)

- Mapping dialogue acts into natural language

inform(name=Seven_Days, foodtype=Chinese)

Seven Days is a nice Chinese restaurant

Template-Based NLG

- Define a set of rules to map frames to NL

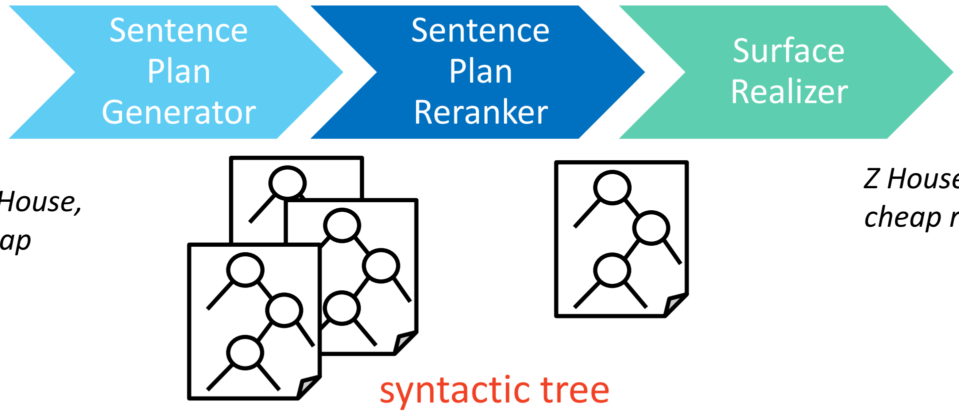
Semantic Frame	Natural Language
confirm()	"Please tell me more about the product your are looking for."
confirm(area=$V)	"Do you want somewhere in the $V?"
confirm(food=$V)	"Do you want a $V restaurant?"
confirm(food=$V,area=$W)	"Do you want a $V restaurant in the $W."

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

Plan-Based NLG (Walker+, 2002)

- Divide the problem into pipeline

Inform(
name=Z_House,
price=cheap
)



Z House is a cheap restaurant.

  - Statistical sentence plan generator (Stent et al., 2009)
  - Statistical surface realizer (Dethlefs et al., 2013; Cuayáhuitl et al., 2014; …)

Pros: can model complex linguistic structures
Cons: heavily engineered, require domain knowledge

Class-Based LM NLG (Oh & Rudnicky, 2000)

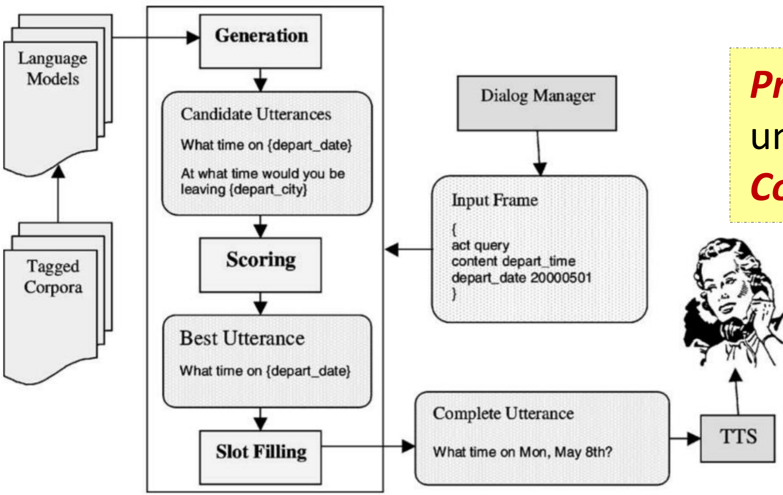
http://dl.acm.org/citation.cfm?id=1117568

- Class-based language modeling

$$P(X \mid c) = \sum_t \log p(x_t \mid x_0, x_1, \cdots, x_{t-1}, c)$$

- NLG by decoding $X^* = \arg\max_X P(X \mid c)$

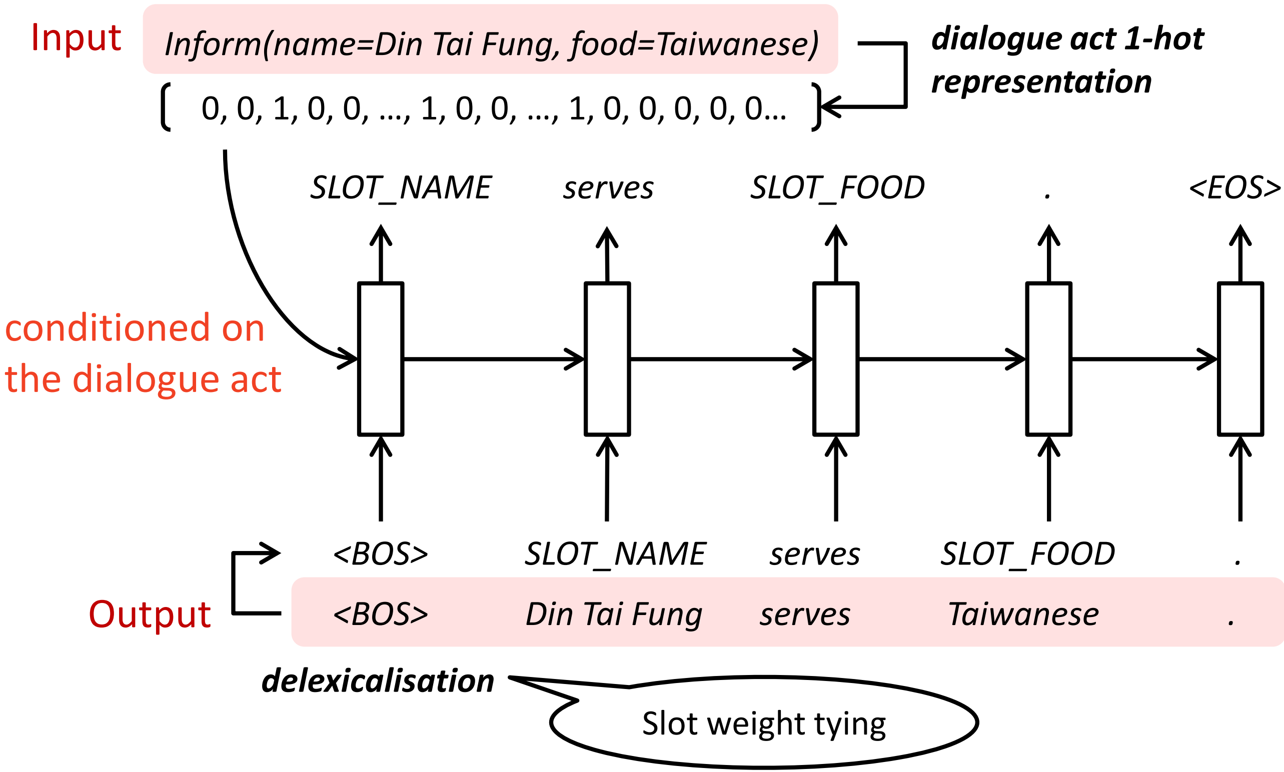
Classes:
inform_area
inform_address
…
request_area
request_postcode



Pros: easy to implement/ understand, simple rules
Cons: computationally inefficient

RNN-Based LM NLG (Wen+, 2015)

http://www.anthology.aclweb.org/W/W15/W15-46.pdf#page=295



Handling Semantic Repetition

- 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.
- Deficiency in either model or decoding (or both)
- Mitigation
  - Post-processing rules (Oh & Rudnicky, 2000)
  - Gating mechanism (Wen et al., 2015)
  - Attention (Mei et al., 2016; Wen et al., 2015)

Semantic Conditioned LSTM (Wen+, 2015)

http://www.aclweb.org/anthology/D/D15/D15-1199.pdf

- Original LSTM cell

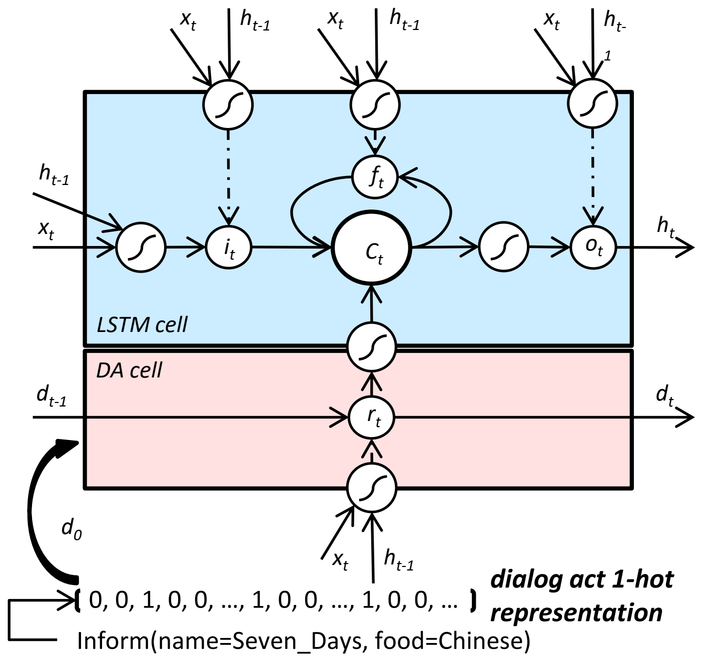
$$\mathbf{i}_t = \sigma(\mathbf{W}_{wi}\mathbf{x}_t + \mathbf{W}_{hi}\mathbf{h}_{t-1})$$
$$\mathbf{f}_t = \sigma(\mathbf{W}_{wf}\mathbf{x}_t + \mathbf{W}_{hf}\mathbf{h}_{t-1})$$
$$\mathbf{o}_t = \sigma(\mathbf{W}_{wo}\mathbf{x}_t + \mathbf{W}_{ho}\mathbf{h}_{t-1})$$
$$\hat{\mathbf{c}}_t = \tanh(\mathbf{W}_{wc}\mathbf{x}_t + \mathbf{W}_{hc}\mathbf{h}_{t-1})$$
$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \hat{\mathbf{c}}_t$$
$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t)$$

- Dialogue act (DA) cell

$$\mathbf{r}_t = \sigma(\mathbf{W}_{wr}\mathbf{x}_t + \mathbf{W}_{hr}\mathbf{h}_{t-1})$$
$$\mathbf{d}_t = \mathbf{r}_t \odot \mathbf{d}_{t-1}$$

- Modify Ct

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \hat{\mathbf{c}}_t + \tanh(\mathbf{W}_{dc}\mathbf{d}_t)$$



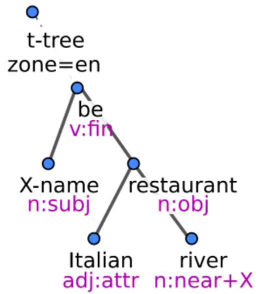
Idea: using gate mechanism to control the generated semantics (dialogue act/slots)

Structural NLG (Dušek & Jurčíček, 2016)

https://www.aclweb.org/anthology/P/P16/P16-2.pdf#page=79

- Goal: NLG based on the syntax tree
  - Encode trees as sequences
  - Seq2Seq model for generation

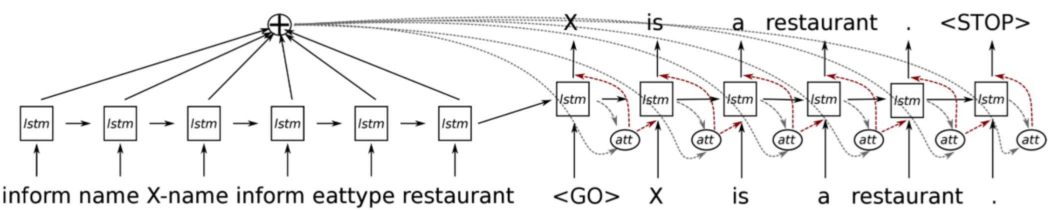
inform(name=X-name,type=placetoeat,eattype=restaurant, area=riverside,food=Italian)



(<root> <root> ((X-name n:subj) be v:fin ((Italian adj:attr) restaurant n:obj (river n:near+X))))

X-name n:subj be v:fin Italian adj:attr restaurant n:obj river n:near+X

X is an Italian restaurant near the river.



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

https://arxiv.org/pdf/1606.03632.pdf; https://ai.google/research/pubs/pub46311

- 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

Structural NLG (Nayak+, 2017)

https://ai.google/research/pubs/pub46311

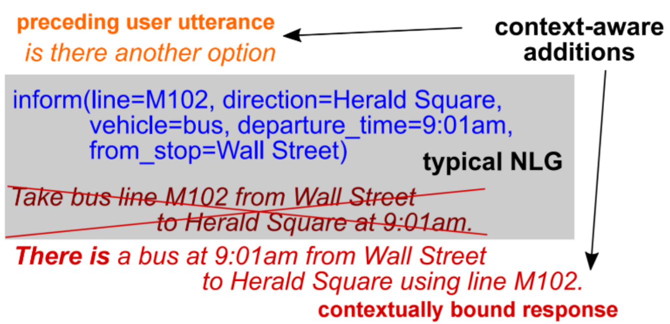
- Sentence plans as part of the input sequence

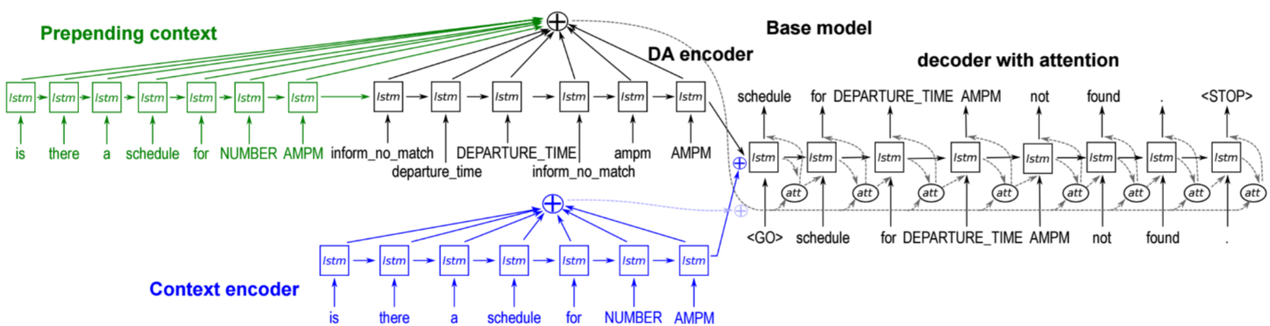
Plan sup.	Input tokens					
NONE	decor	decent	service	decent	quality	good
FLAT	decor	decent	service	decent		
	quality	good				
POSITIONAL	<B>	decor	decent	service	decent	
	<I>	quality	good			

Contextual NLG (Dušek & Jurčíček, 2016)

https://www.aclweb.org/anthology/W/W16/W16-36.pdf#page=203

- Can we do better by using context ?
- Goal: provide context-aware responses
  - Context encoder
  - Seq2Seq model

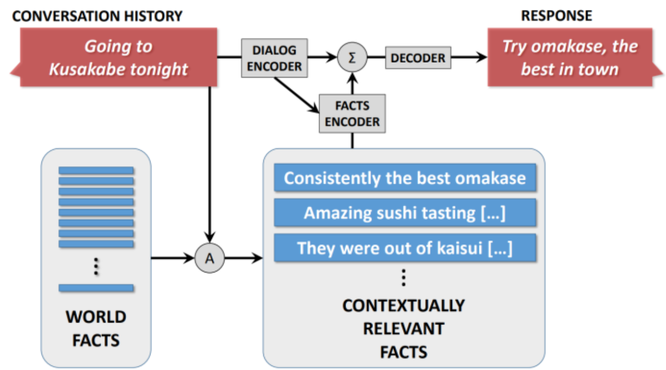




Knowledge-Grounded Conversation Model

(Ghazvininejad+, 2017)

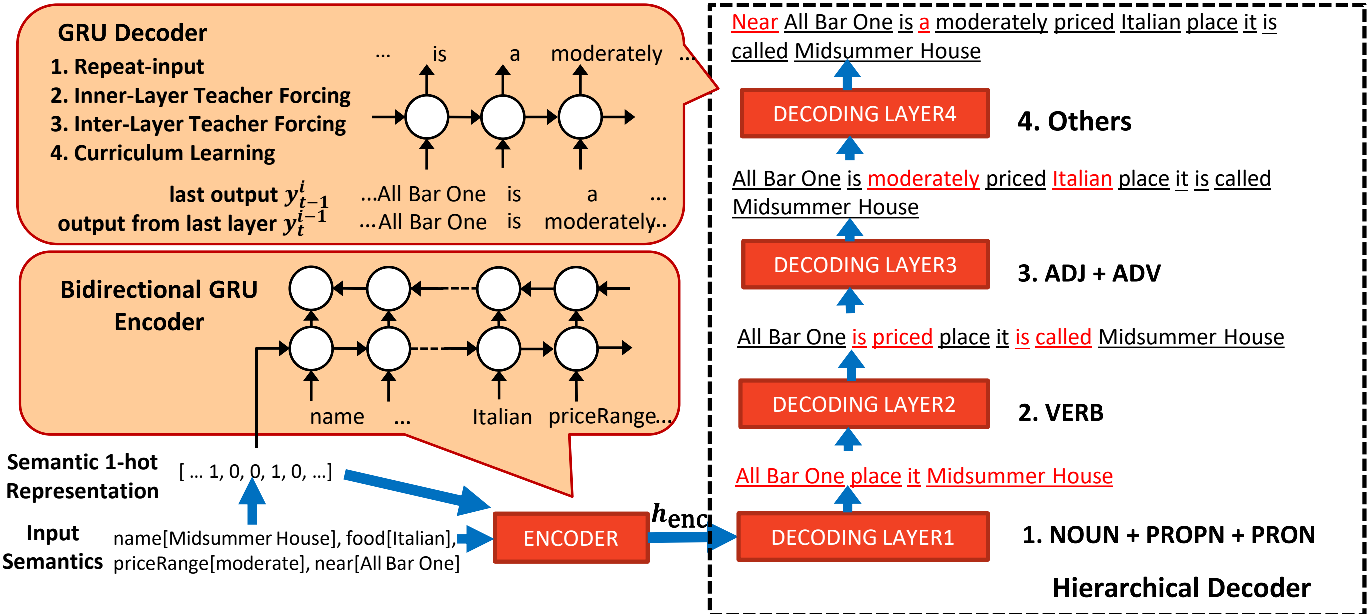
User input: *Going to Kusakabe tonight.*
Neural model: Have a great time!
Human: You'll love it! Try omasake, the best in town.



A: **Looking forward to trying** *@pizzalibretto* **tonight! my expectations are high**. B: **Get** *the rocco salad*. **Can you eat** *calamari*?	
A: **Anyone in** *Chi* **have a** *dentist office* **they recommend? I'm never going back to** *[...]* **and would love a reco!** B: **Really looved** *Ora* **in** *Wicker Park*.	
A: **I'm at** *California Academy of Sciences* B: **Make sure you catch** *the show* **at** *the Planetarium*. **Tickets are usually limited.**	
A: **I'm at** *New Wave Cafe*. B: **Try to get to** *Dmitri's* **for dinner**. **Their** *pan fried scallops* **and** *shrimp scampi* **are to die for.**	
A: **I just bought:** *[...] 4.3-inch portable GPS navigator* **for my wife, shh, don't tell her**. B: **I heard this brand** *loses battery power*.	

Hierarchical NLG w/ Linguistic Patterns (Su+, 2018)

https://arxiv.org/abs/1808.02747

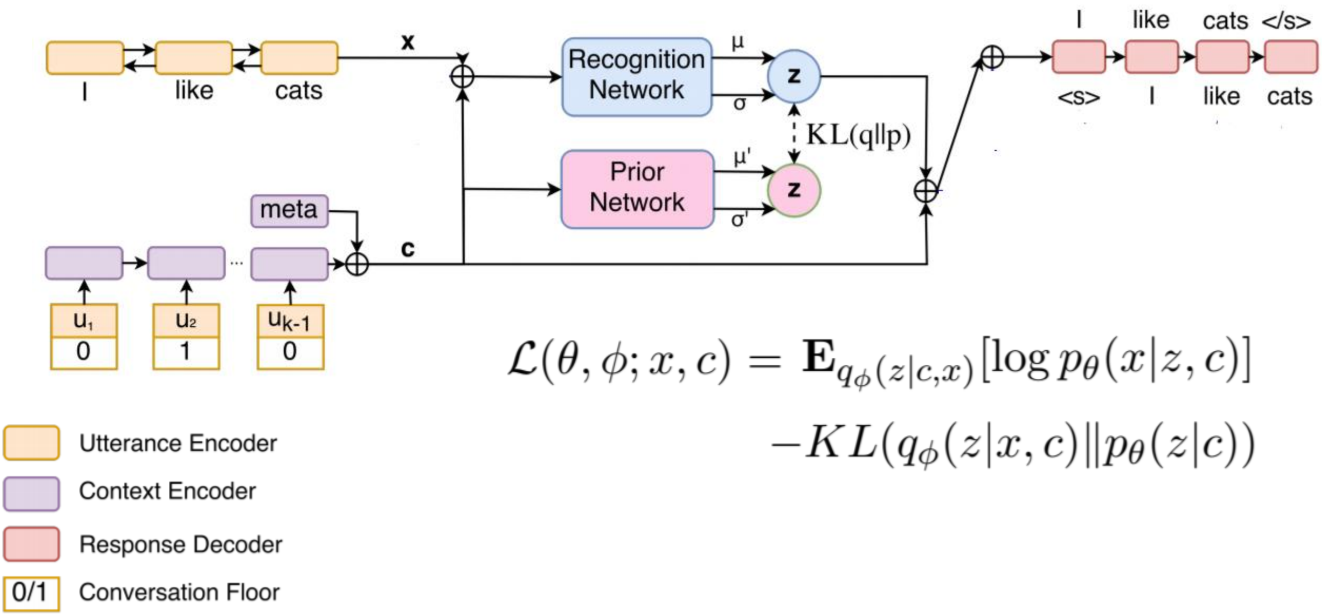


Idea: gradually generate words based on the linguistic knowledge

Learning Discourse-Level Diversity (Zhao+, 2017)

http://aclweb.org/anthology/P/P17/P17-1061.pdf

- Conditional VAE
- Improves diversity of responses

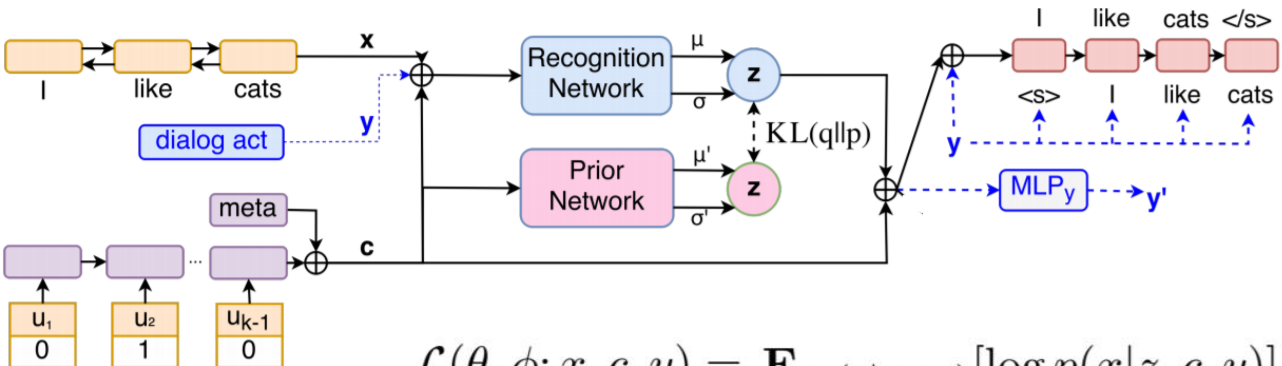


$$\mathcal{L}(\theta, \phi; x, c) = \mathbf{E}_{q_\phi(z|c,x)}[\log p_\theta(x|z,c)] \\ -KL(q_\phi(z|x,c) \| p_\theta(z|c))$$

Learning Discourse-Level Diversity (Zhao+, 2017)

http://aclweb.org/anthology/P/P17/P17-1061.pdf

- Conditional VAE
- Improves diversity of responses with dialogue acts

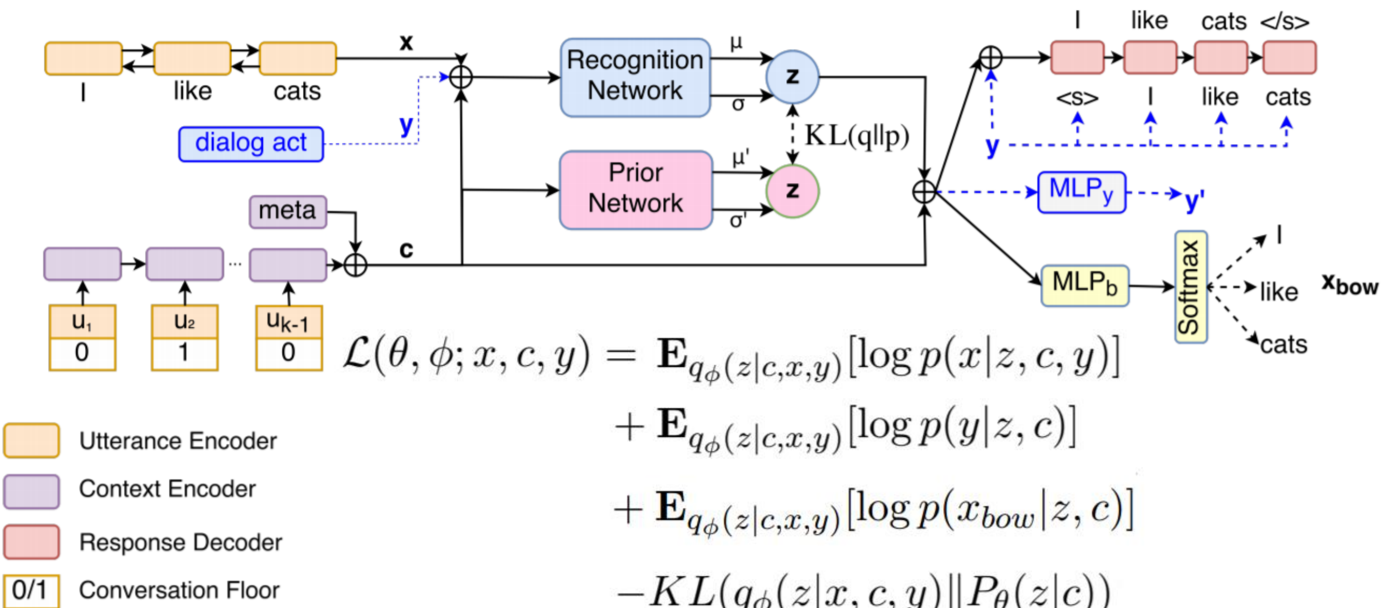


$$\mathcal{L}(\theta,\phi;x,c,y) = \mathbf{E}_{q_\phi(z|c,x,y)}[\log p(x|z,c,y)] + \mathbf{E}_{q_\phi(z|c,x,y)}[\log p(y|z,c)] - KL(q_\phi(z|x,c,y)\|P_\theta(z|c))$$

Learning Discourse-Level Diversity (Zhao+, 2017)

http://aclweb.org/anthology/P/P17/P17-1061.pdf

- Knowledge guided conditional VAE
- Improves diversity of responses with dialogue acts



$$\mathcal{L}(\theta,\phi;x,c,y) = \mathbf{E}_{q_\phi(z|c,x,y)}[\log p(x|z,c,y)] + \mathbf{E}_{q_\phi(z|c,x,y)}[\log p(y|z,c)] + \mathbf{E}_{q_\phi(z|c,x,y)}[\log p(x_{bow}|z,c)] - KL(q_\phi(z|x,c,y)\|P_\theta(z|c))$$

NLG Evaluation

- Metrics
  - Subjective: human judgement (Stent+, 2005)
    - Adequacy: correct meaning
    - Fluency: linguistic fluency
    - Readability: fluency in the dialogue context
    - Variation: multiple realizations for the same concept
  - Objective: automatic metrics
    - Word overlap: BLEU (Papineni+, 2002), METEOR, ROUGE
    - Word embedding based: vector extrema, greedy matching, embedding average

There is a gap between human perception and automatic metrics

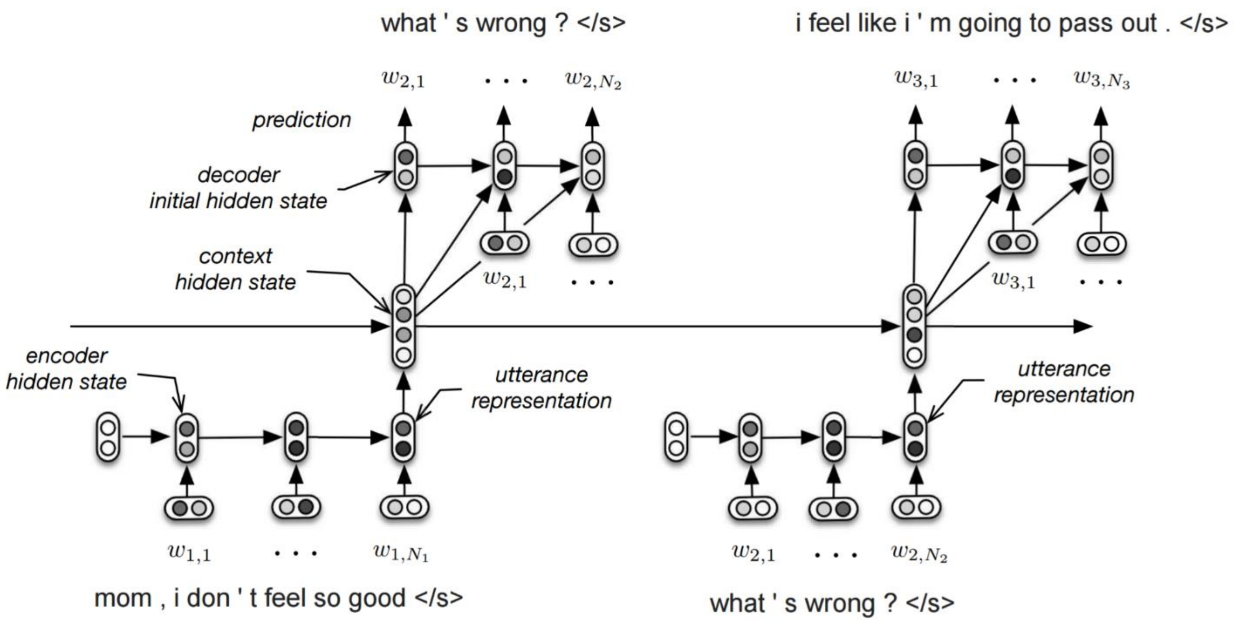
Outline

- Introduction & Background
  - Neural Networks
  - Reinforcement Learning
- Modular Dialogue System
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management (DM)
    - Dialogue State Tracking (DST)
    - Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - **End-to-End Neural Dialogue Systems**
- System Evaluation
- Recent Trends on Learning Dialogues

ChitChat Hierarchical Seq2Seq (Serban+, 2016)

http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/11957

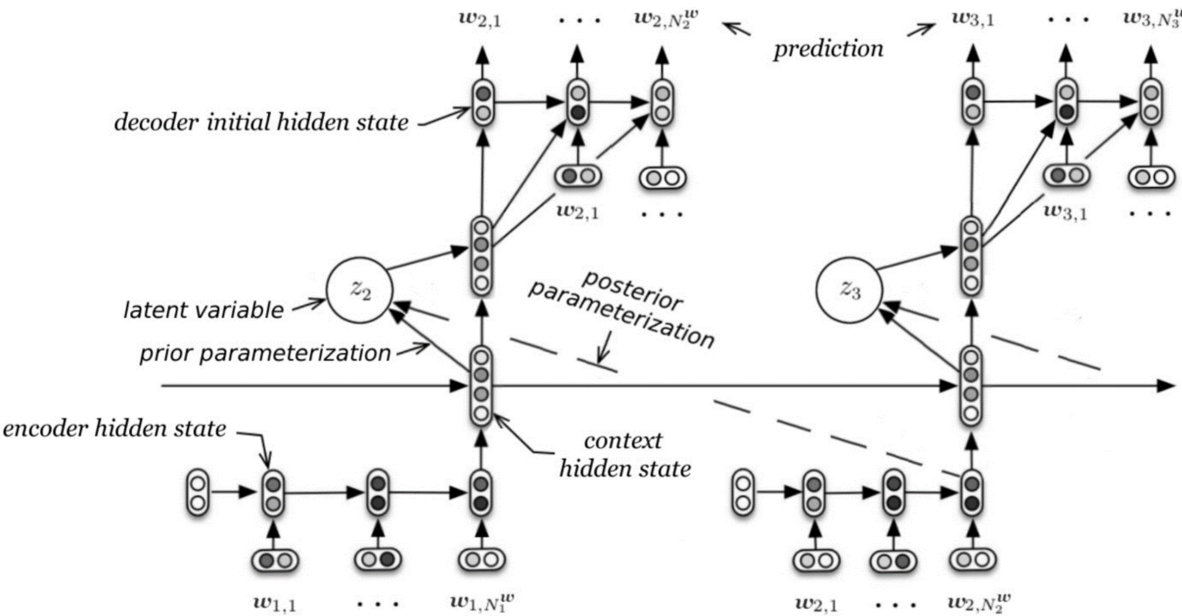
- A hierarchical seq2seq model for generating dialogues



ChitChat Hierarchical Seq2Seq (Serban+, 2017)

https://arxiv.org/abs/1605.06069

- A hierarchical seq2seq model with Gaussian latent variable for generating dialogues

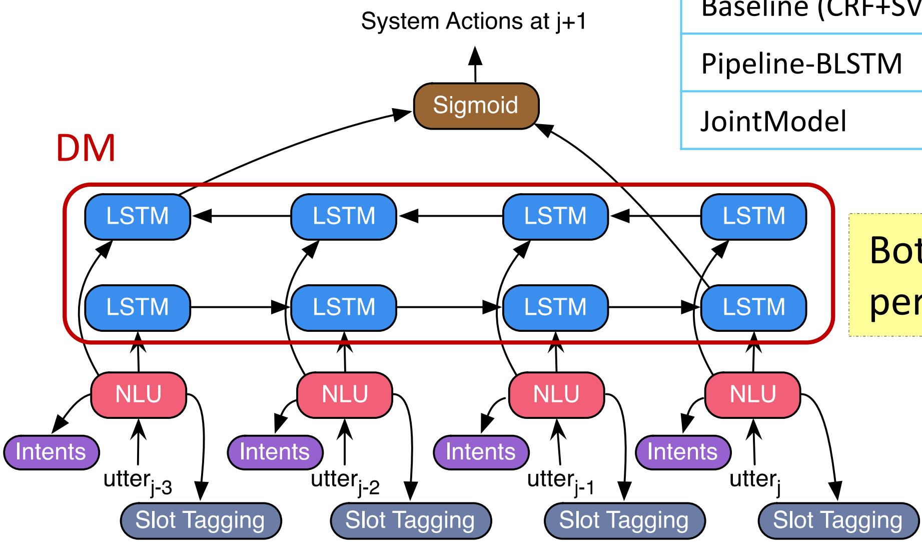


E2E Joint NLU and DM (Yang+, 2017)

https://arxiv.org/abs/1612.00913

- Idea: errors from DM can be propagated to NLU for better robustness

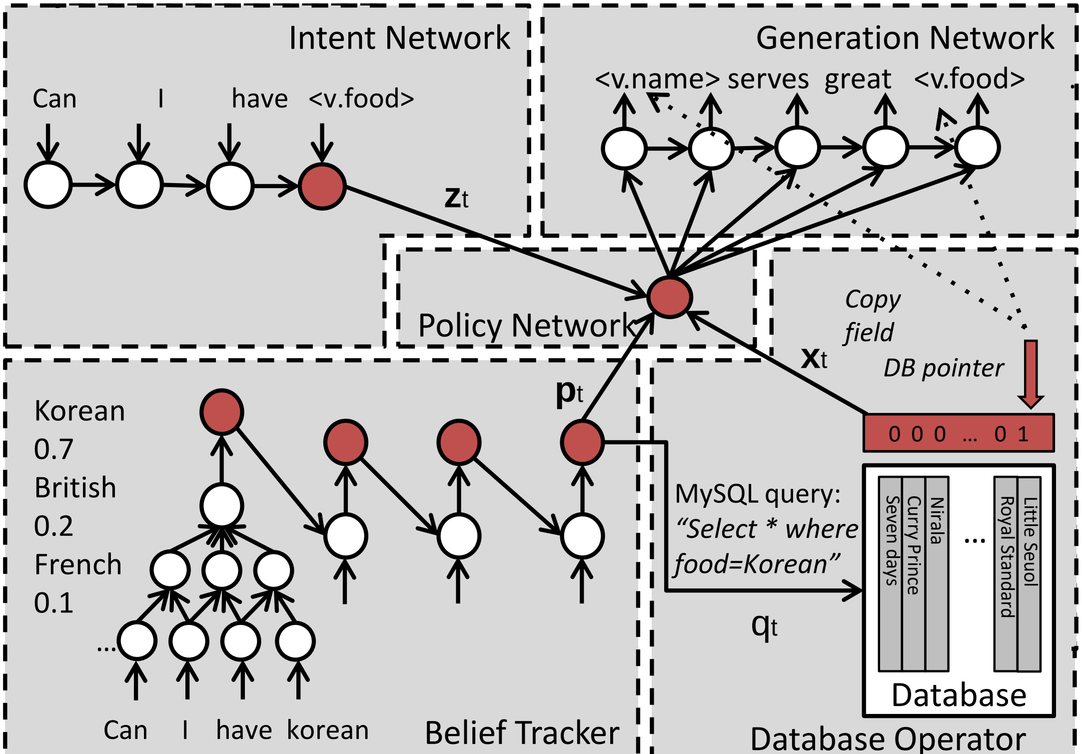
Model	DM	NLU
Baseline (CRF+SVMs)	7.7	33.1
Pipeline-BLSTM	12.0	36.4
JointModel	**22.8**	**37.4**



Both DM and NLU performance is improved

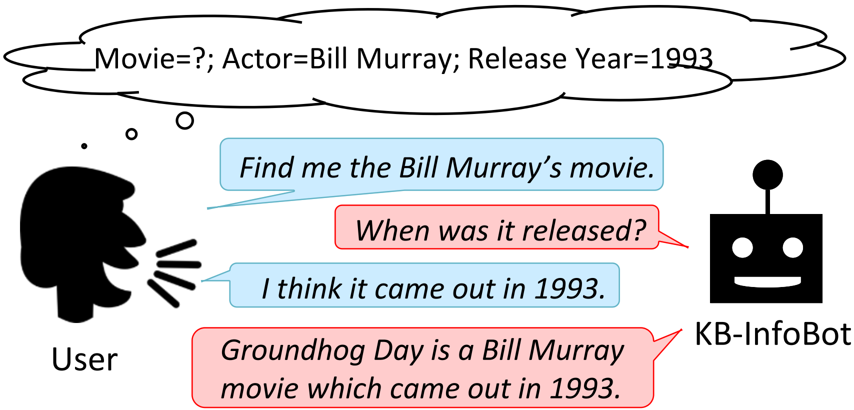
E2E Supervised Dialogue System (Wen+, 2016)

https://arxiv.org/abs/1604.04562



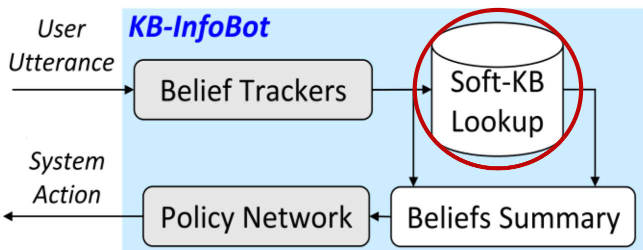
E2E RL-Based Info-Bot (Dhingra+, 2016)

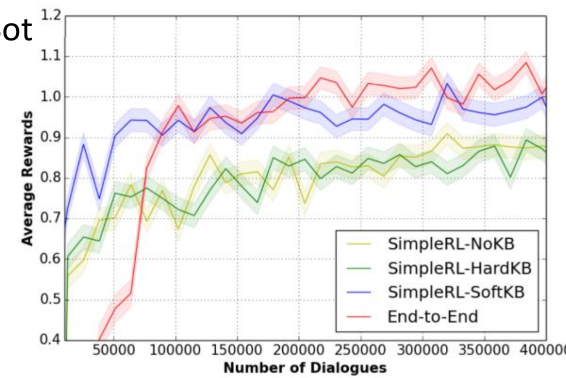
http://www.aclweb.org/anthology/P/P17/P17-1045.pdf



Entity-Centric Knowledge Base

Movie	Actor	Year
Groundhog Day	Bill Murray	1993
Australia	Nicole Kidman	X
Mad Max: Fury Road	X	2015





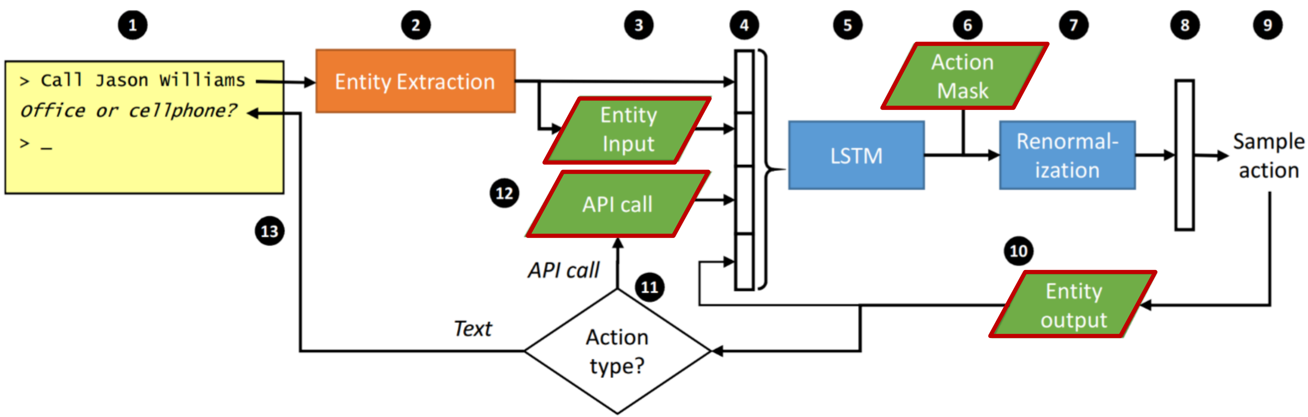
Idea: differentiable database for propagating the gradients

E2E LSTM-Based Dialogue Control

(Williams & Zweig, 2016)

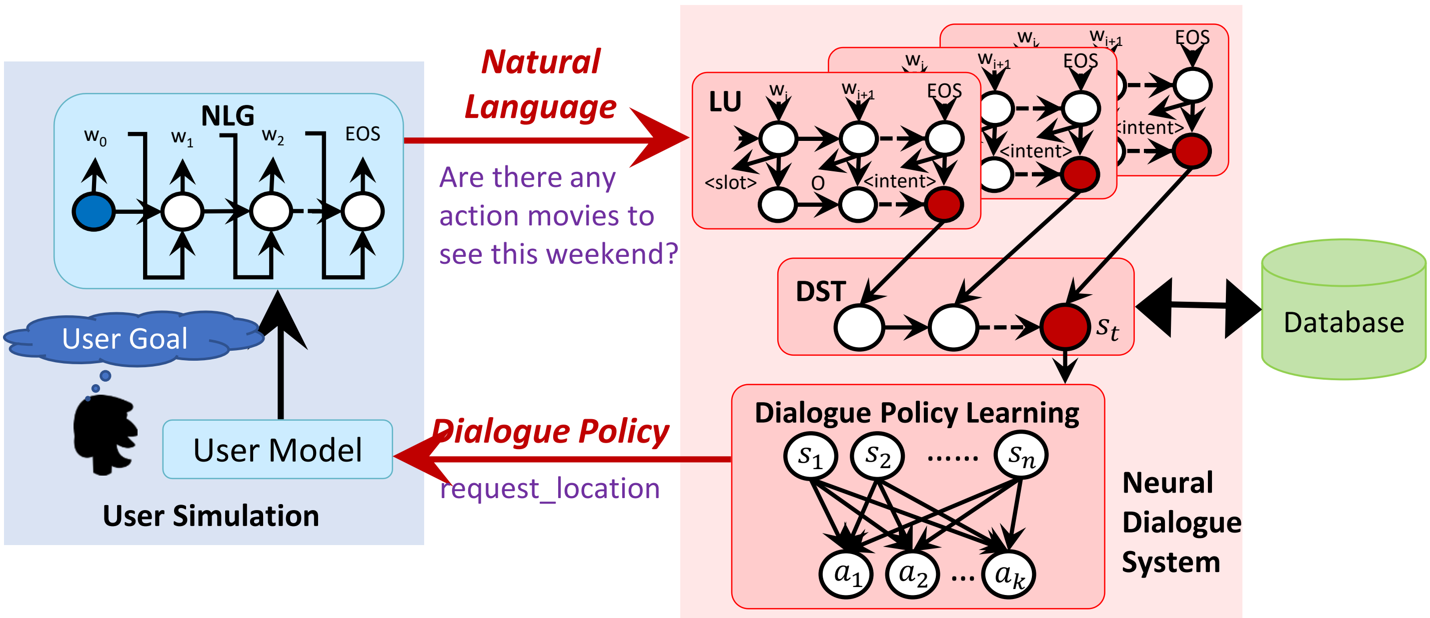
https://arxiv.org/abs/1606.01269

- Idea: an LSTM maps from raw dialogue history directly to a distribution over system actions
  - Developers can provide software including business rules & programmatic APIs
    → LSTM can take actions in the real world on behalf of the user
  - The LSTM can be optimized using SL or RL



E2E Task-Completion Bot (TC-Bot) (Li+, 2017)

http://www.aclweb.org/anthology/I17-1074



Idea: SL for each component and RL for end-to-end training

RL TC-Bot (Li+, 2017)

http://www.aclweb.org/anthology/I17-1074

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, thanks.
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 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

E2E Imitation and RL Agent (Liu+, 2018)

http://aclweb.org/anthology/N18-1187

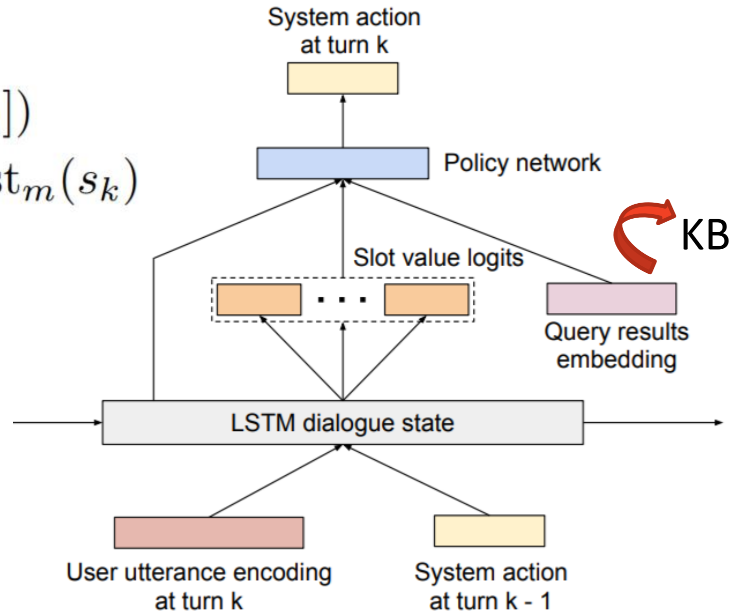
- Generate distribution over candidate slot values:

$$s_k = \mathrm{LSTM}(s_{k-1},\ [U_k,\ A_{k-1}])$$
$$P(l_k^m \mid \mathbf{U}_{\leq k},\ \mathbf{A}_{<k}) = \mathrm{SlotDist}_m(s_k)$$

- Generate system action:

$$P(a_k \mid U_{\leq k},\ A_{<k},\ E_{\leq k})$$
$$= \mathrm{PolicyNet}(s_k, v_k, E_k)$$

- Train Supervised → REINFORCE



Dialogue Challenge

- DSTC: Dialog System Technology Challenge

Challenge	Track	Theme
DSTC6	Track 1	End-to-End Goal-Oriented Dialog Learning
	Track 2	End-to-End Conversation Modeling
	Track 3	Dialogue Breakdown Detection
DSTC7	Track 1	Sentence Selection
	Track 2	Sentence Generation
	Track 3	AVSD: **A**udio **V**isual **S**cene-aware **D**ialog

- SLT 2018 Microsoft Dialogue Challenge:
  End-to-End Task-Completion Dialogue Systems
- The Conversation Intelligence Challenge:
  ConvAI2 - PersonaChat

System Evaluation

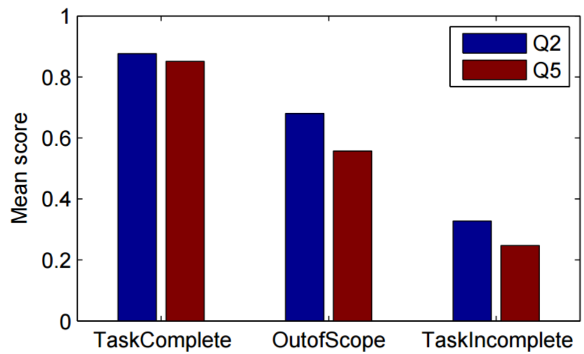
Dialogue System Evaluation

- Dialogue model evaluation
  - Crowd sourcing
  - User simulator
- Response generator evaluation
  - Word overlap metrics
  - Embedding based metrics

Crowdsourcing for Evaluation (Yang+, 2012)

http://www-scf.usc.edu/~zhaojuny/docs/SDSchapter_final.pdf

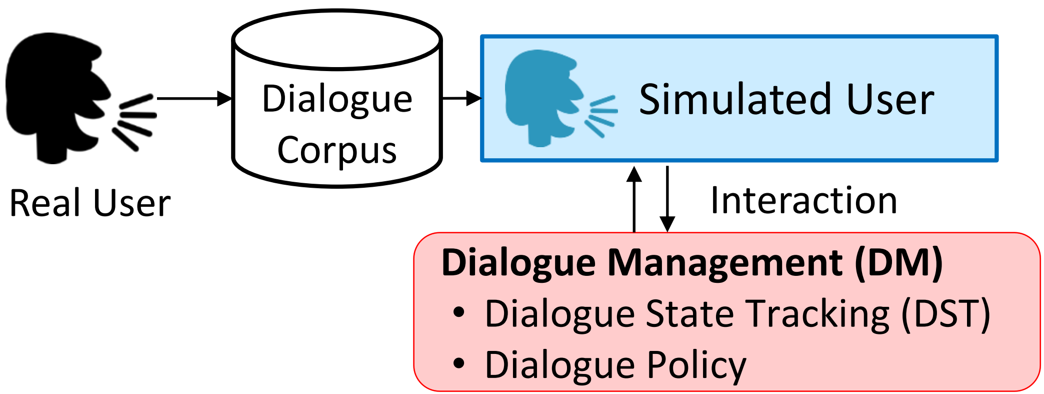
Q1	**Do you think you understand from the dialog what the user wanted?**
Opt	1) No clue 2) A little bit 3) Somewhat 4) Mostly 5) Entirely
Aim	*elicit the Worker's confidence in his/her ratings.*
Q2	**Do you think the system is successful in providing the information that the user wanted?**
Opt	1) Entirely unsuccessful 2) Mostly unsuccessful 3) Half successful/unsuccessful 4) Mostly successful 5) Entirely successful
Aim	*elicit the Worker's perception of whether the dialog has fulfilled the informational goal of the user.*
Q3	**Does the system work the way you expect it?**
Opt	1) Not at all 2) Barely 3) Somewhat 4) Almost 5) Completely
Aim	*elicit the Worker's impression of whether the dialog flow suits general expectations.*
Q4	**Overall, do you think that this is a good system?**
Opt	1) Very poor 2) Poor 3) Fair 4) Good 5) Very good
Aim	*elicit the Worker's overall impression of the SDS.*
Q5	**What category do you think the dialog belongs to?**
Opt	1) Task is incomplete 2) Out of scope 3) Task is complete
Aim	*elicit the Worker's impression of whether the dialog reflects task completion.*



The normalized mean scores of Q2 and Q5 for approved ratings in each category. A higher score maps to a higher level of task success

User Simulation

- Goal: Generate natural and reasonable conversations to enable reinforcement learning for exploring the policy space
- Conventional corpora cannot be used to train RL agents.
- Simulator is replaced by crowd users to replicate real environment.



Material: http://deepdialogue.miulab.tw

keeps a list of its goals and actions → randomly generates an agenda → updates its list of goals and adds new ones

User Simulation

- First, generate a user goal.
- The user goal contains:
  - Dialog act
  - Inform slots
  - Request slots

start-time="4 pm"

date="today"

city="Birmingham"

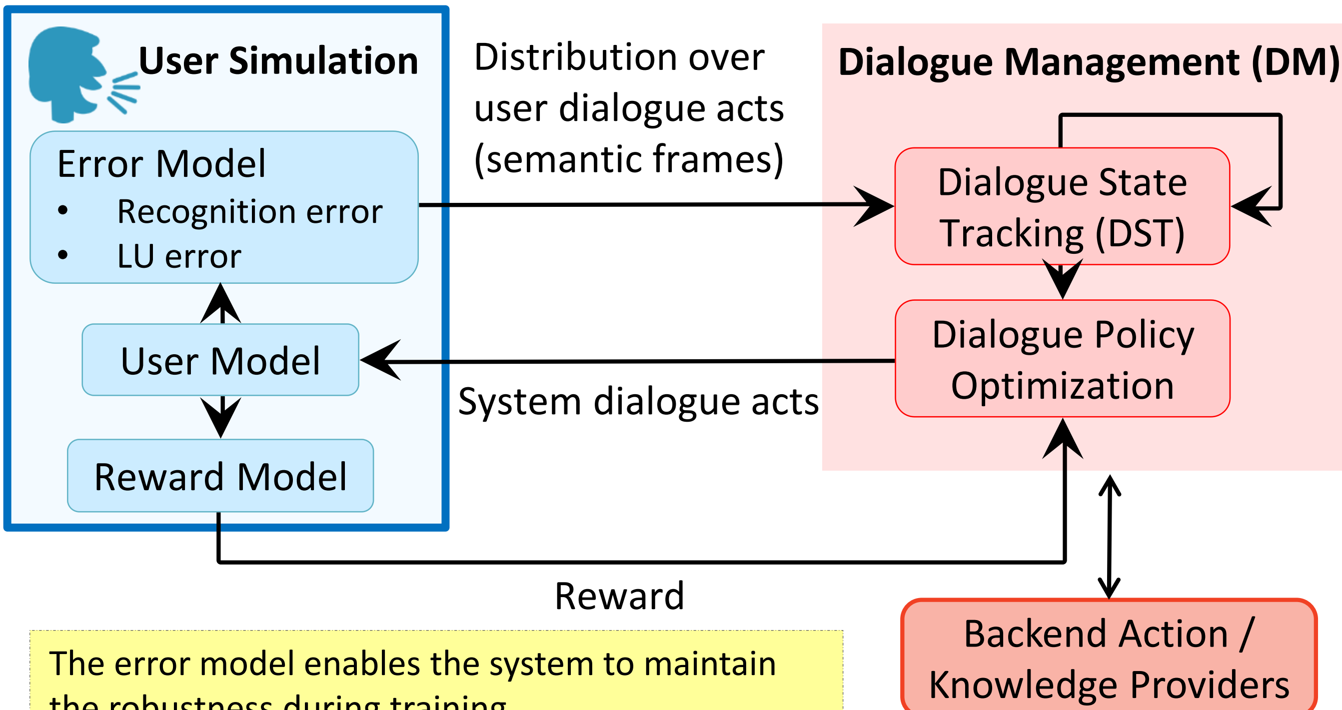
Are there any tickets available for 4 pm ?

'Hidden Figures' is playing at 4pm and 6 pm.

What is playing in Birmingham theaters today ?


```
{
  "request_slots": {
    "ticket": "UNK",
    "theater": "UNK"
  },
  "diaact": "request",
  "inform_slots": {
    "city": "birmingham",
    "numberofpeople": "2",
    "state": "al",
    "starttime": "4 pm",
    "date": "today",
    "moviename": "deadpool"
  }
}
```

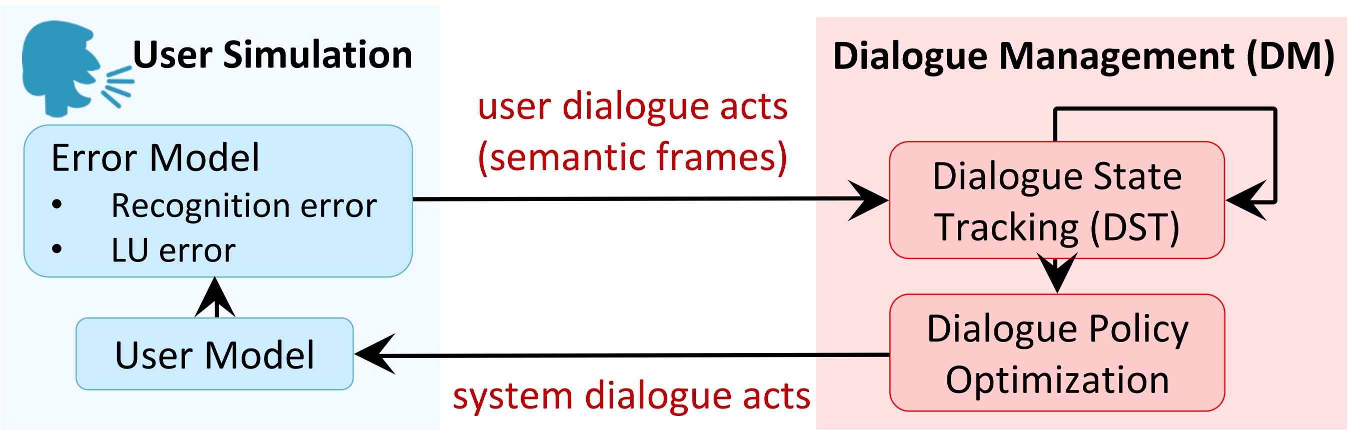
Elements of User Simulation



The error model enables the system to maintain the robustness during training

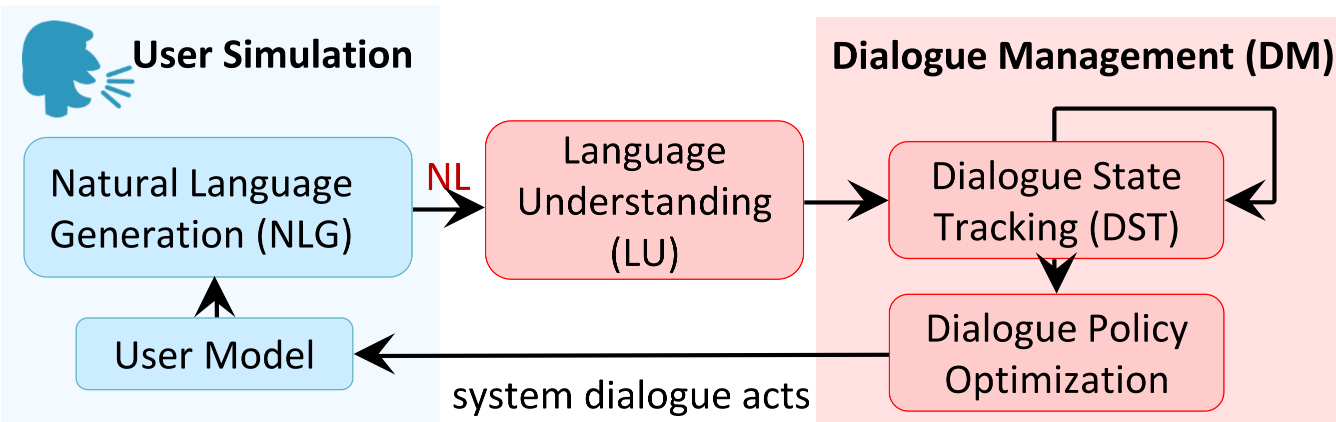
Frame-Level Interaction

- DM receives frame-level information
  - No error model: perfect recognizer and LU
  - Error model: simulate the possible errors



Natural Language Level Interaction

- User simulator sends natural language
  - No recognition error
  - Errors from NLG or LU



Rule-Based Simulator for RL Agent (Li+, 2016)

http://arxiv.org/abs/1612.05688

- rule-based simulator + collected data
- starts with sets of goals, actions, KB, slot types
- publicly available simulation framework
- movie-booking domain: ticket booking and movie seeking
- provide procedures to add and test own agent

```
class AgentDQN(Agent):
    def run_policy(self, representation):
        """ epsilon-greedy policy """

        if random.random() < self.epsilon:
            return random.randint(0, self.num_actions - 1)
        else:
            if self.warm_start == 1:
                if len(self.experience_replay_pool) > self.experience_replay_pool_size:
                    self.warm_start = 2
                return self.rule_policy()
            else:
                return self.dqn.predict(representation, {}, predict_model=True)

    def train(self, batch_size=1, num_batches=100):
        """ Train DQN with experience replay """

        for iter_batch in range(num_batches):
            self.cur_bellman_err = 0
            for iter in range(len(self.experience_replay_pool)/(batch_size)):
                batch = [random.choice(self.experience_replay_pool) for i in xrange(batch_size)]
                batch_struct = self.dqn.singleBatch(batch, {'gamma': self.gamma}, self.clone_dqn)
```

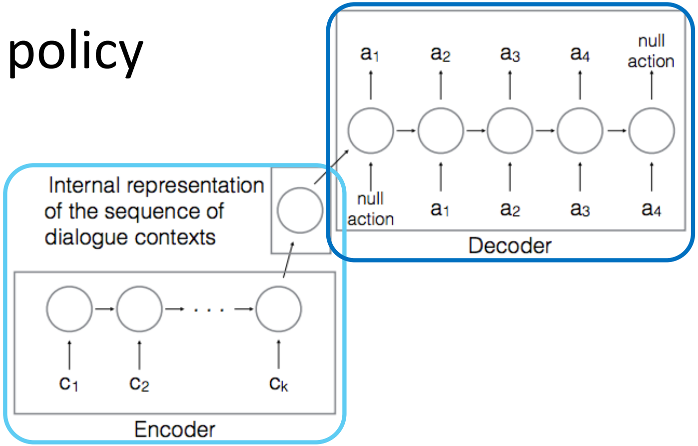
Model-Based User Simulators

- Bi-gram models (Levin+, 2000)
- Graph-based models (Scheffler & Young, 2000)
- Data Driven Simulator (Jung+, 2009)
- Neural Models (deep encoder-decoder)

Seq2Seq User Simulation (El Asri+, 2016)

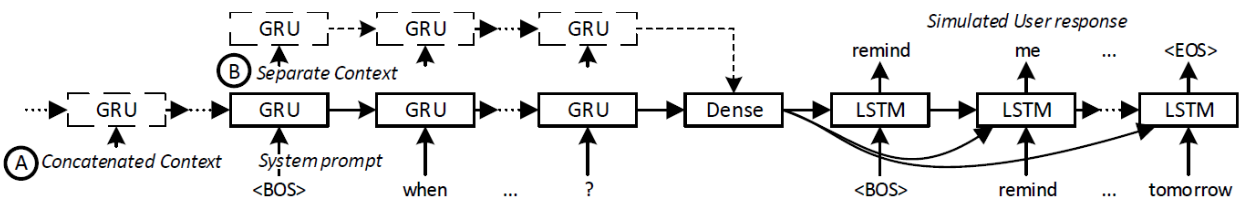
https://arxiv.org/abs/1607.00070

- Seq2Seq trained from dialogue data
  - Input: c_i encodes contextual features, such as the previous system action, consistency between user goal and machine provided values
  - Output: a dialogue act sequence form the user
- Extrinsic evaluation for policy



Seq2Seq User Simulation (Crook & Marin, 2017)

- Seq2Seq trained from dialogue data
  - No labeled data
  - Trained on just human to machine conversations



User Simulator for Dialogue Evaluation

Understanding Ability

- whether **constrained values** specified by users can be understood by the system
- agreement percentage of system/user understandings over the entire dialog (averaging all turns)

Efficiency

- Number of dialogue turns
- Dissimilarity between the dialogue turns (larger is better)

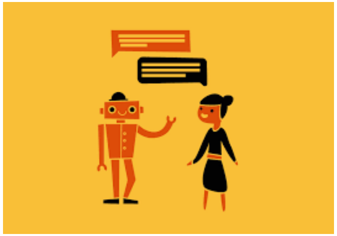
Action Appropriateness

- an explicit **confirmation** for an uncertain user utterance is an appropriate system action
- providing information based on misunderstood user requirements

How NOT to Evaluate Dialogue System (Liu+, 2017)

https://arxiv.org/pdf/1603.08023.pdf

- How to evaluate the quality of the generated response ?
  - Specifically investigated for chat-bots
  - Crucial for task-oriented tasks as well
- Metrics:
  - Word overlap metrics, e.g., BLEU, METEOR, ROUGE, etc.
  - Embeddings based metrics, e.g., contextual/meaning representation between target and candidate



Dialogue Response Evaluation (Lowe+, 2017)

- Problems of existing automatic evaluation
  - can be biased
  - correlate poorly with human judgements of response quality
  - using word overlap may be misleading
- Solution
  - collect a dataset of accurate human scores for variety of dialogue responses (e.g., coherent/un-coherent, relevant/irrelevant, etc.)
  - use this dataset to train an automatic dialogue evaluation model – learn to compare the reference to candidate responses!
  - Use RNN to predict scores by comparing against human scores!

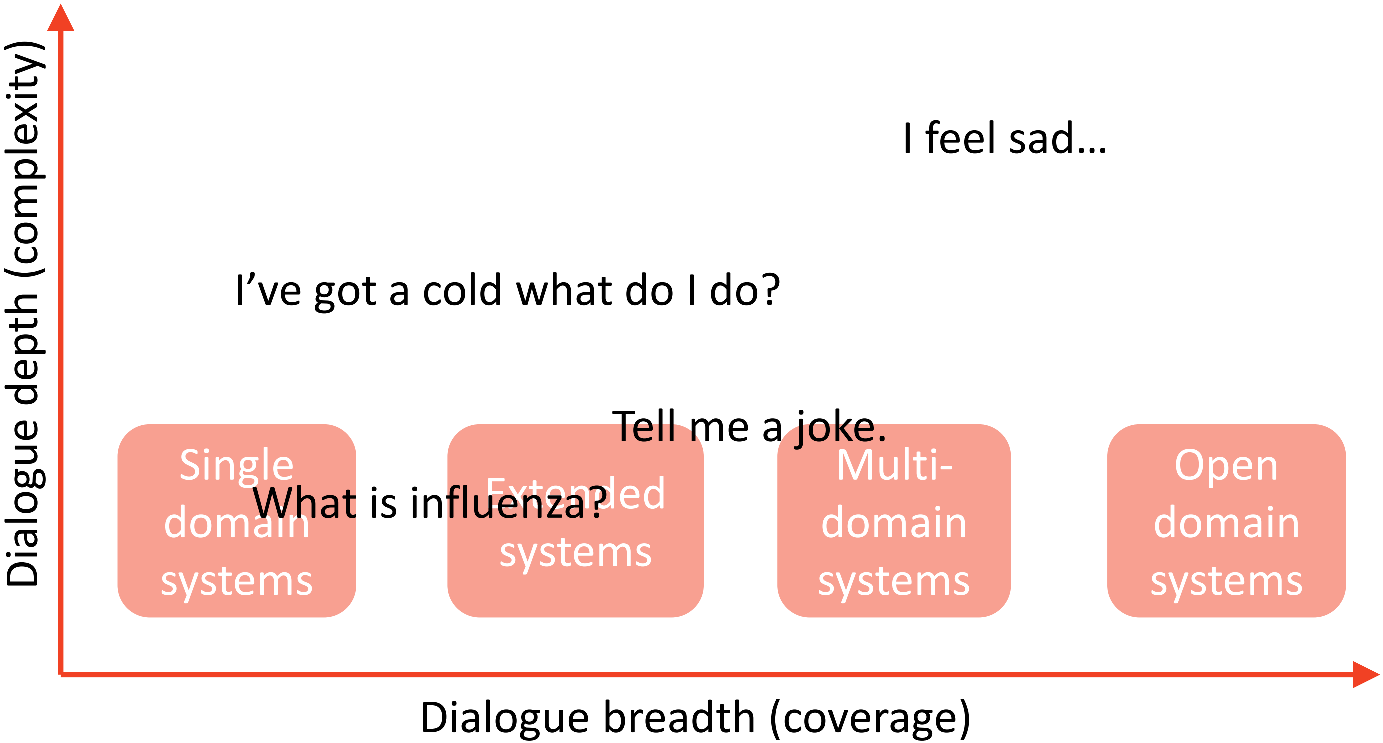
Context of Conversation **Speaker A:** *Hey, what do you want to do tonight?* **Speaker B:** *Why don't we go see a movie?*
Model Response *Nah, let's do something active.*
Reference Response *Yeah, the film about Turing looks great!*

Recent Trends and Challenges

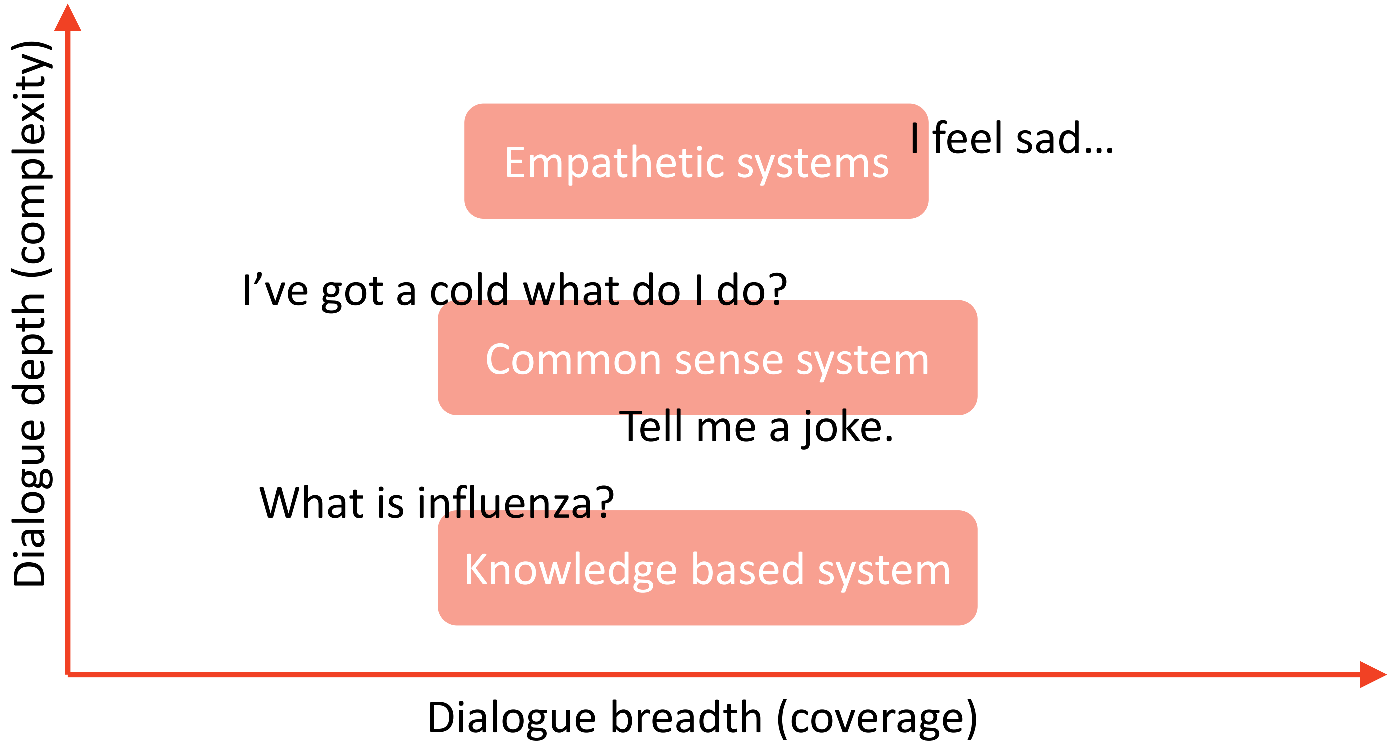
Dialogue Breadth

Dialogue Depth

Evolution Roadmap



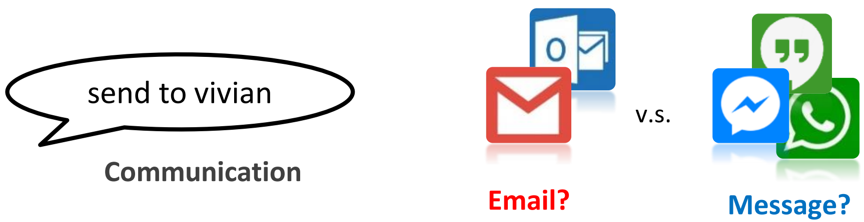
Evolution Roadmap



App Behavior for Understanding (Chen+, 2015)

http://dl.acm.org/citation.cfm?id=2820781

- Task: user intent prediction
- Challenge: language ambiguity



① **User preference**
- ✓ Some people prefer "Message" to "Email"
- ✓ Some people prefer "Ping" to "Text"

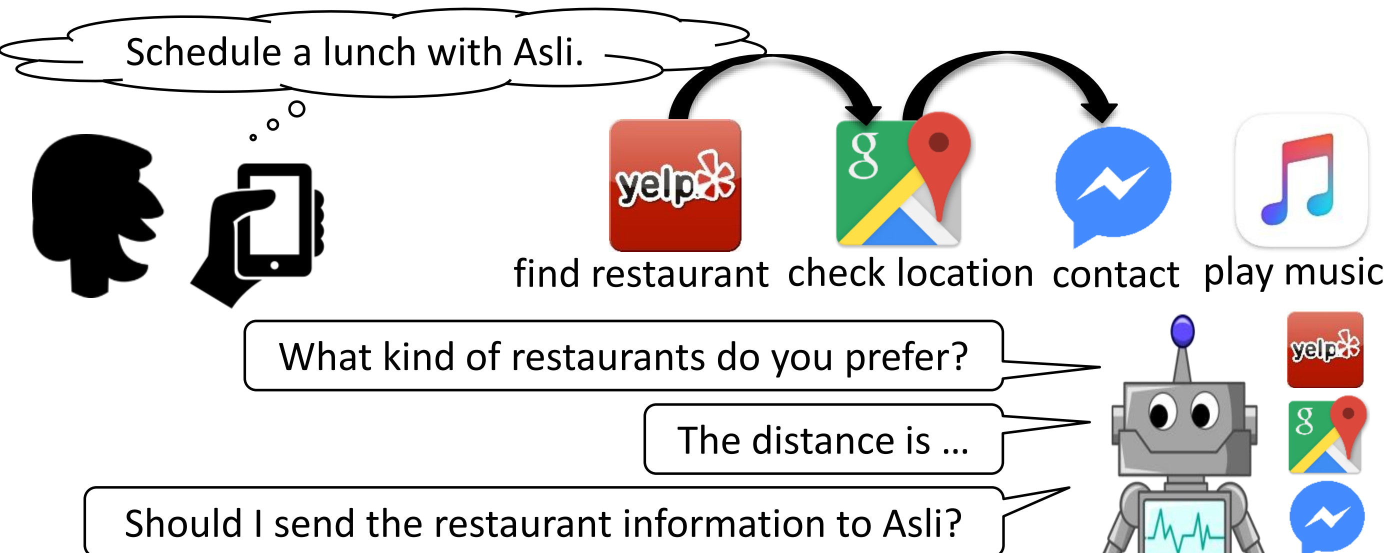
② **App-level contexts**
- ✓ "Message" is more likely to follow "Camera"
- ✓ "Email" is more likely to follow "Excel"

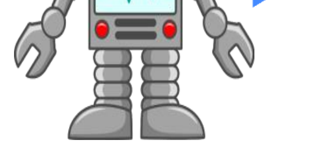
Considering behavioral patterns in history to model understanding for intent prediction.

High-Level Intention for Dialogue Planning
(Sun+, 2016; Sun+, 2016)

http://dl.acm.org/citation.cfm?id=2856818; http://www.lrec-conf.org/proceedings/lrec2016/pdf/75_Paper.pdf

- High-level intention may span several domains





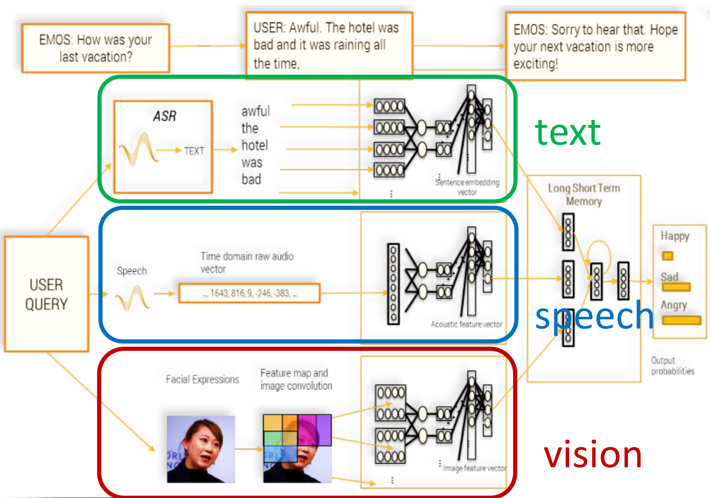
Users can interact via <u>high-level descriptions</u> and the system learns <u>how to plan the dialogues</u>

Empathy in Dialogue System (Fung+, 2016)

https://arxiv.org/abs/1605.04072

- Embed an empathy module
  - Recognize emotion using multimodality
  - Generate emotion-aware responses

Zara - The Empathetic Supergirl



Emotion Recognizer

Made with love by Ivo Technologies in collaboration with Hong Kong University of Science and Technology

Face recognition output

(index):1728
(index):1729
{
"recognition": "Race: Asian Confidence: 65.42750000000001 Smiling: 3.95896 Gender: Female Confidence: 88.9369",
"race": "Asian",
"race_confidence": "65.42750000000001",
"smiling": "3.95896",
"gender": "Female",
"gender_confidence": "88.9369"
}

Visual Object Discovery through Dialogues
(Vries+, 2017)

https://arxiv.org/pdf/1611.08481.pdf

- Recognize objects using “Guess What?” game
- Includes “spatial”, “visual”, “object taxonomy” and “interaction”

Is it a person?	*No*
Is it an item being worn or held?	*Yes*
Is it a snowboard?	*Yes*
Is it the red one?	*No*
Is it the one being held by the person in blue?	*Yes*

Is it a cow?	*Yes*
Is it the big cow in the middle?	*No*
Is the cow on the left?	*No*
On the right ?	*Yes*
First cow near us?	*Yes*

Conclusion

Summary of Challenges

- HCI is a hot topic but modelling requires dealing with several moving components
- Most state-of-the-art models are based on DNN:
  - Requires a lot of labeled and unlabeled data
  - Hyperparameter-tunning
- Reasoning and Interpretability:
  - To reduce bias
  - Improve performance

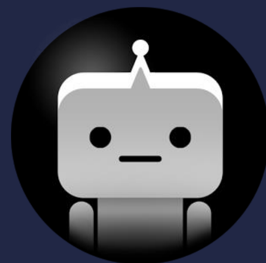
Summary of Challenges

- Data collection is usually from unstructured data
- Most systems have complex cascaded architectures:
  - Requires high accuracy
  - Careful hyperparameter tuning
  - Correct objective function setting/balancing

Brief Conclusions

- Introduce recent deep learning methods used in dialogue models
- Highlight main components of dialogue systems and new deep learning architectures used for these components
- Talk about challenges and new avenues for current state-of-the-art research
- Provide all materials online!

http://deepdialogue.miulab.tw



Thanks to Tsung-Hsien Wen, Pei-Hao Su, Li Deng, Jianfeng Gao, Sungjin Lee, Milica Gašić, Lihong Li, Xiujin Li, Abhinav Rastogi, Ankur Bapna, Pararth Shah, Shyam Udaphyay and Gokhan Tur for sharing their slides.

THANKS FOR YOUR ATTENTION!