ICML | 2019

Thirty-sixth International Conference on Machine Learning

Neural Approaches to Conversational AI

Jianfeng Gao, Michel Galley

Microsoft Research

ICML 2019

Long Beach, June 10, 2019

Book details:

https://www.nowpublishers.com/article/Details/INR-074

https://arxiv.org/abs/1809.08267 (preprint)

Contact Information:

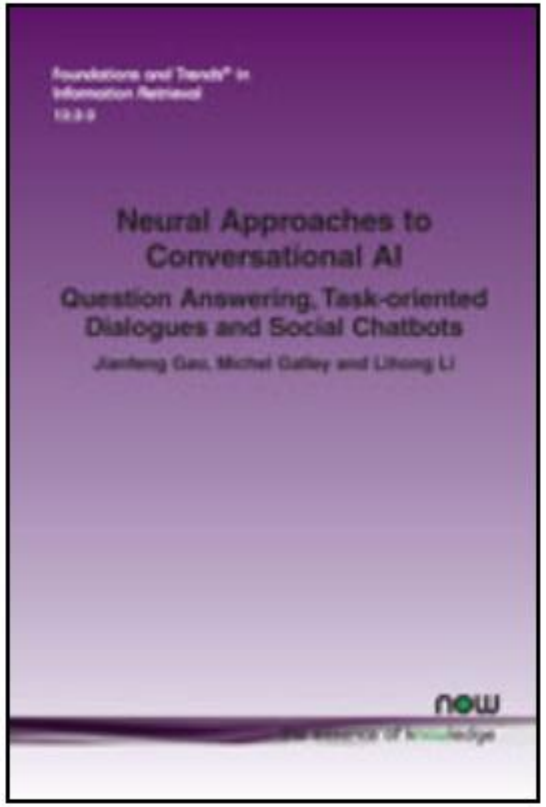
Jianfeng Gao http://research.microsoft.com/~jfgao

Michel Galley http://research.microsoft.com/~mgalley

Slides:

http://microsoft.com/en-us/research/publication/neural-approaches-to-conversational-ai/

We thank **Lihong Li, Bill Dolan** and **Yun-Nung (Vivian) Chen** for contributing slides.



Outline

- **Part 1: Introduction**
  - **Who should attend this tutorial**
  - **Dialogue: what kinds of problem**
  - **A unified view: dialogue as optimal decision making**
  - **Deep learning leads to paradigm shift in NLP**
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogues
- Part 4: Fully data-driven conversation models and chatbots

Who should attend this tutorial?

- Whoever wants to understand and create modern dialogue agents that
  - Can chat like a human (to establish long-term emotional connections with users)
  - Can answer questions of various topics (movie stars, theory of relativity)
  - Can fulfill tasks (whether report, travel planning)
  - Can help make business decision
  - …
- Focus on neural approaches, but hybrid approaches are widely used.

Aspirational Goal: Enterprise Assistant

Where are sales lagging behind our forecast?

Task Completion

The worst region is [country], where sales are XX% below projections

Do you know why?

QA (decision support)

The forecast for [product] growth was overly optimistic

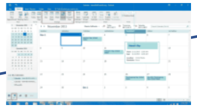
How can we turn this around?

Here are the 10 customers in [country] with the most growth potential, per our CRM model

Can you set up a meeting with the CTO of [company]?

Info Consumption

Task Completion



Yes, I’ve set up a meeting with [person name] for next month when you’re in [location]

Thanks

What kinds of problems?

“I am smart”	Turing Test (“I” talk like a human)
“I have a question”	Information consumption
“I need to get this done”	Task completion
“What should I do?”	Decision support

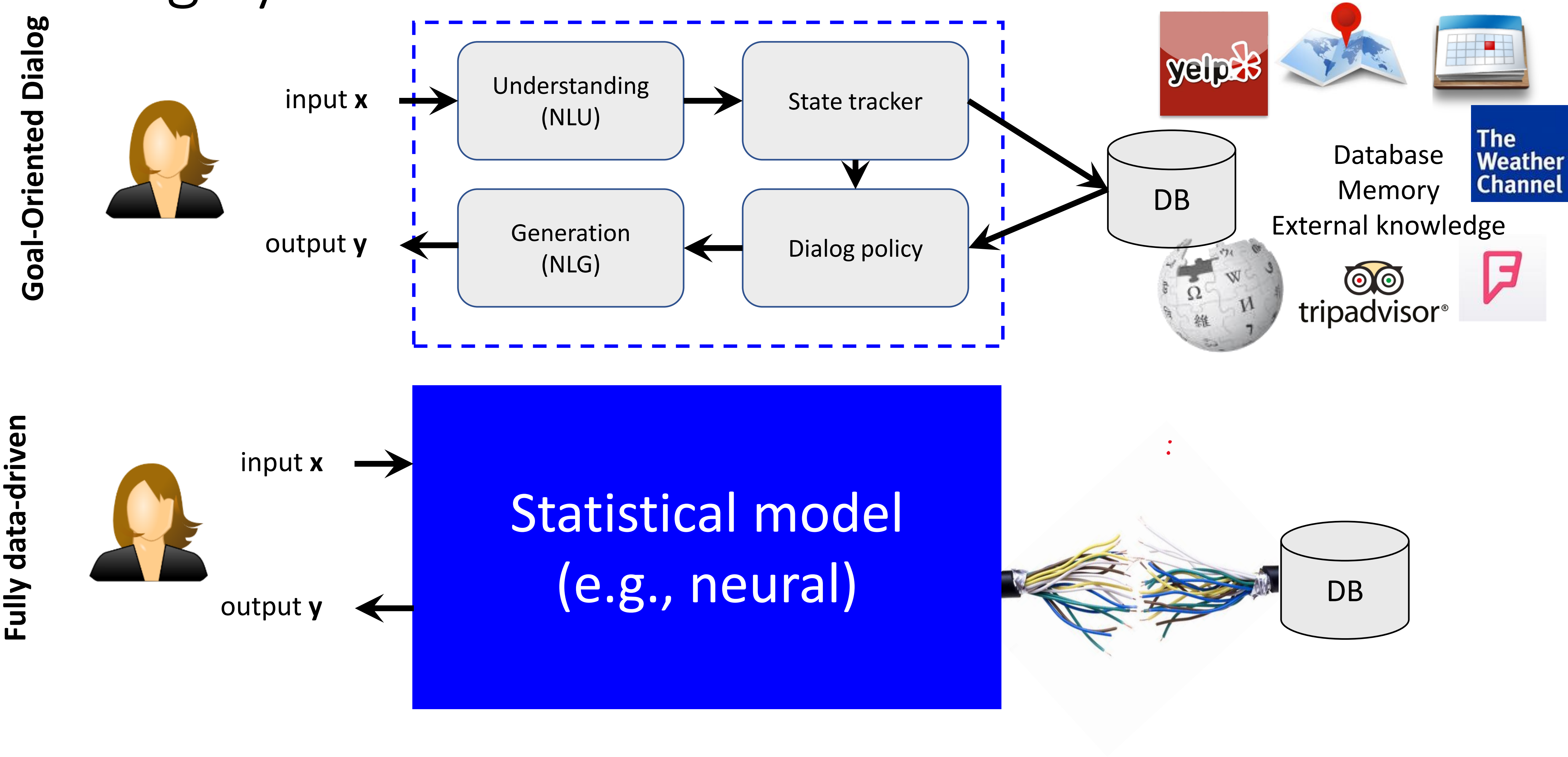
What kinds of problems?

Chitchat (social bot)

“I am smart”	Turing Test (“I” talk like a human)
“I have a question”	Information consumption
“I need to get this done”	Task completion
“What should I do?”	Decision support

Goal-oriented dialogues

Dialog Systems



[Young+ 13; Tur & De Mori 11; Ritter+ 11; Sordoni+ 15; Vinyals & Le 15; Shang+ 15; etc.]

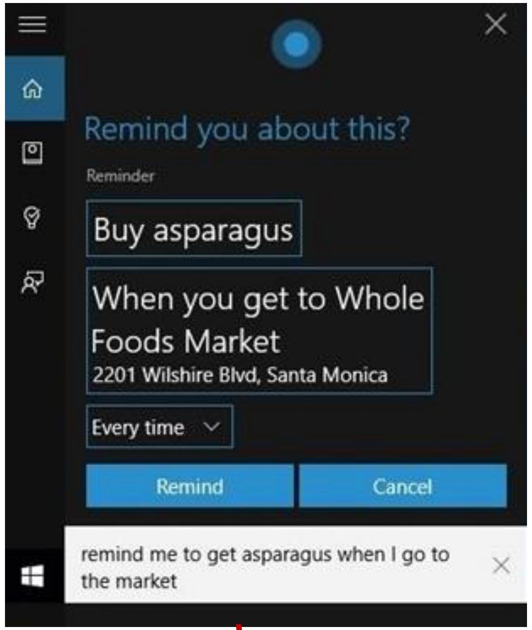
A unified view: dialogue as optimal decision making

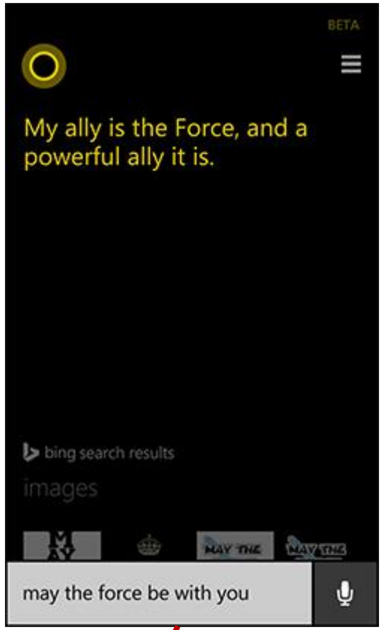
- Dialogue as a Markov Decision Process (MDP)
  - Given state s, select action a according to (hierarchical) policy π
  - Receive reward r, observe new state s'
  - Continue the cycle until the episode terminates.
- Goal of dialogue learning: find optimal π to maximize expected rewards

Dialogue	State (s)	Action (a)	Reward (r)
Info Bots (Q&A bot over KB, Web etc.)	Understanding of user Intent (belief state)	Clarification questions, Answers	Relevance of answer # of turns (less is better)
Task Completion Bots (Movies, Restaurants, …)	Understanding of user goal (belief state)	Dialog act + slot_value	Task success rate # of turns (less is better)
Social Bot (XiaoIce)	Conversation history	Response	Engagement, # of turns (more is better)

Personal assistants today

Google Now Siri Cortana amazon echo

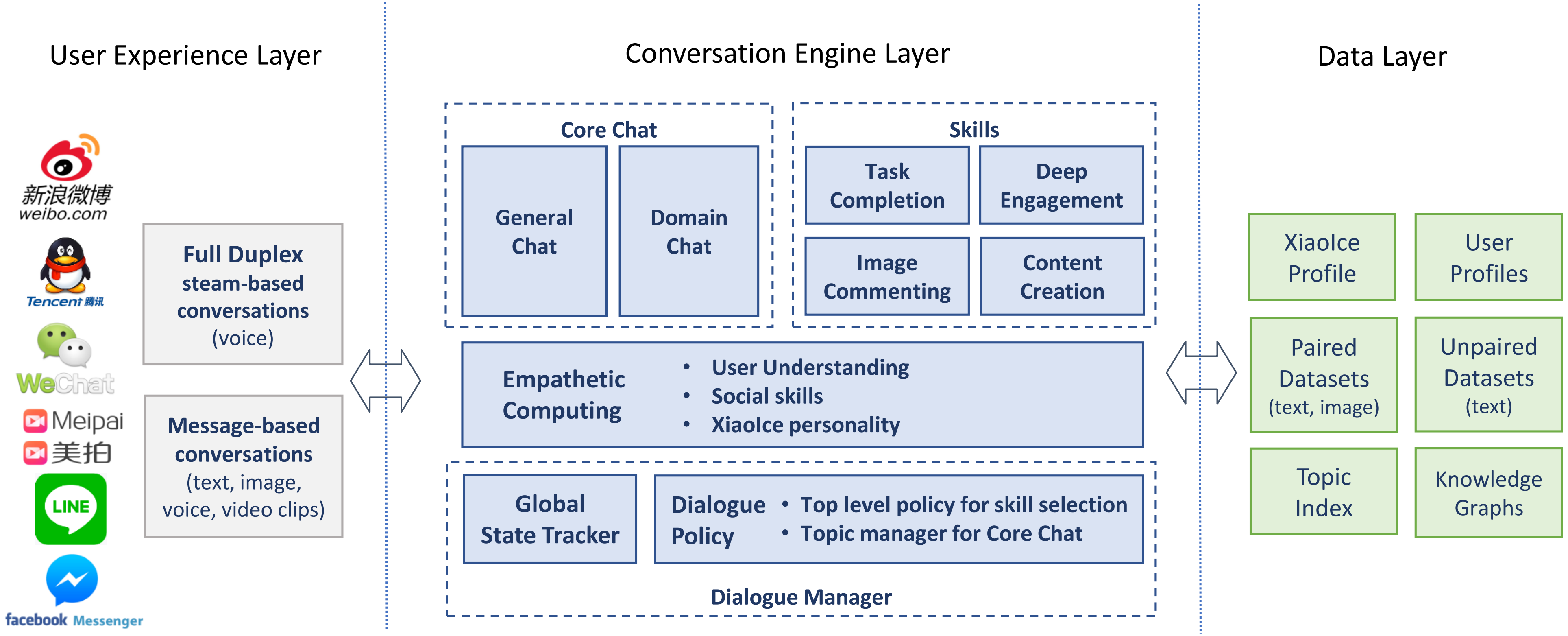




goal oriented

Engaging (social bots)

XiaoIce System Architecture

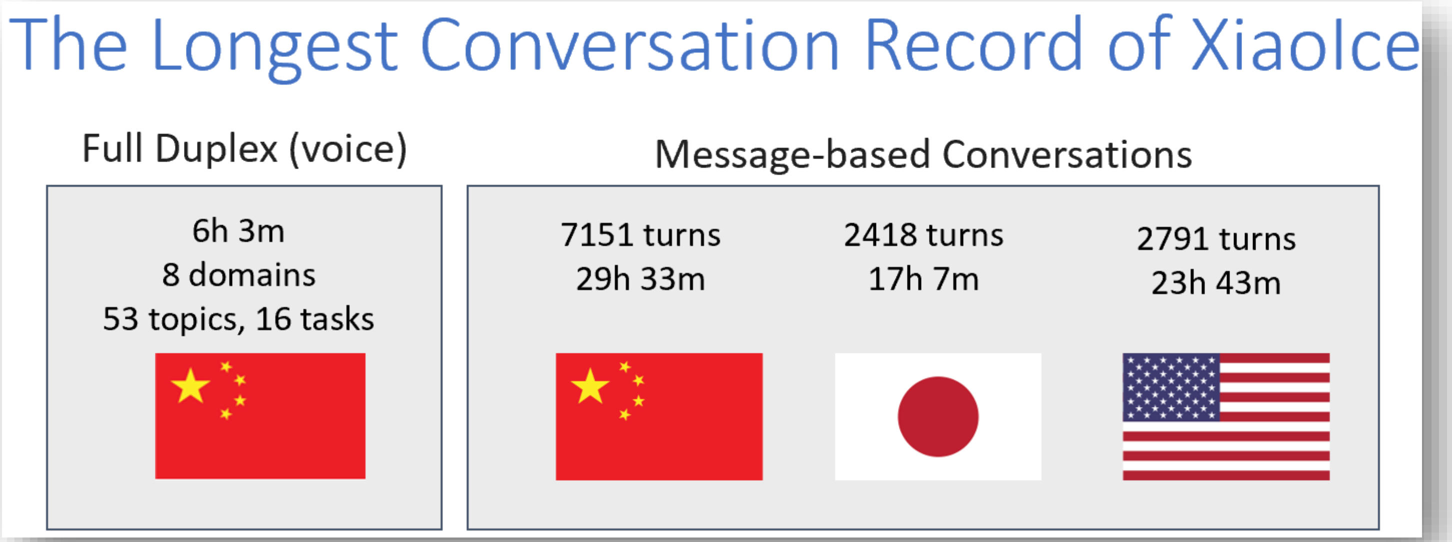


[Design and Implementation of XiaoIce, an empathetic social chatbot]

General Chat Skill	1	[Human] Hey, I am so nervous	[人类]哎！我好紧张
		[Xiaoice] What’s the matter? Why are you so nervous?	[小冰]什么事，这么不淡定?
	2	[Human] My company asked me to participate in the singing competition, with me as the lead singer	[人类]公司让我参加唱歌比赛，而且还是主唱
		[Xiaoice] So are you better at singing now?	[小冰]那你唱歌水平提高了?
	3	[Human] Anyway, at average level of my company	[人类]反正是我们公司的水平
Music Chat Skill		[Xiaoice] Not everyone can be a lead singer. You are so great!	[小冰]可不是谁都能当主唱的真棒!
	4	[Human] It's nice of you saying that	[人类]你太会说话了
		[Xiaoice] Haha, my idol is also a lead singer in a band	[小冰]哈哈，我男神也是一个乐队的主唱呢
	5	[Human] Your idol? Who’s that?	[人类]你男神？谁啊
		[Xiaoice] Ashin from Mayday	[小冰]五月天的阿信啊
	6	[Human] Umm...Okay	[人类]呃。。。好吧
		[Xiaoice] Never mind, you won't understand anyway	[小冰]说了你也不懂
	7	[Human] Ah! Boring	[人类]哎。。郁闷
		[Xiaoice] What is it?	[小冰]怎么了?
	8	[Human] Nothing, just being cute	[人类]没事，只是日常撒撒娇
		[Xiaoice] Give you a hug. Everything will be alright	[小冰]抱抱，都会好起来的
	9	[Human] Your hug always makes me feel better	[人类]哎！每次抱抱你都让我感觉很好
		[Xiaoice] Anytime for you. A nice sleep will help you relax	[小冰]嗯，好好睡一觉是最管用的
	10	[Xiaoice] Okay, I am going to bed	[人类]好那我睡了
		[Xiaoice] Go ahead, I'm going to see my idol in dream too	[小冰]去吧，我也要去梦里见我男神了
	11	[Human] You like Ashin that much, huh?	[人类]你这么喜欢阿信啊
		[Xiaoice] Why not? You don't like him?	[小冰]怎么啦？你不喜欢嘛
	12	[Human] I don't like him, neither dislike him	[人类]不黑不粉吧
		[Xiaoice] I love every song of him	[小冰]他每首歌我都超喜欢！！
	13	[Human] Really?	[人类]真的假的。。
Switch to a new topic		[Xiaoice] Especially *The Time Machine*	[小冰]尤其是那首《时光机》。。
	14	[Human] Oh? Never heard of that	[人类]哦？没听过
		[Xiaoice]Shall I send you the song?	[小冰]发你一起听听?
	15	[Human] Okay	[人类]好啊
		[Xiaoice] Don't know if you will like it or not~	[小冰]不知道你会不会喜欢～
Song-On-Demand Skill		[Xiaoice] [Play the music Card]	[小冰][音乐播放Card]
	16	[Human] That is really touching. I will collect it as my favorite	[人类]有点感动了。。收藏了
		[Xiaoice] That is specially for you. How can I be so nice	[小冰]专门转给你的。。我怎么这么好
	17	[Human] You are always very nice	[人类]你就是很好啊
		[Xiaoice] Hahaha	[小冰]嘿嘿嘿
Ticket-Booking Skill	18	[Human] This is really a good song. Any concert of him coming up recently?	[人类]这首歌真心很好听，他最近有演唱会吗
		[Xiaoice] No. Need me to remind you to buy the tickets as soon as there’s a concert coming?	[小冰]没，要我到时候第一时间通知你买票吗?
	19	[Human] Sure	[人类]好的呀
		[Xiaoice] Okay, will do	[小冰]OK! 我记下啦

XiaoIce: the Most Popular Social Chatbot in the World [Zhou+ 18]

- 660 million users worldwide
- 5 countries: China, Japan, USA, India, Indonesia
- 40 platforms, e.g., WeChat, QQ, Weibo, FB Messenger, LINE
- Average CPS of 23 (better than human conversations)



Session 1

	English	Chinese
1	[Human] Hello, Xiaoice.	[人类]小冰你好
	[Xiaoice] Hello, your profile pic is interesting. Like I've seen it somewhere.	[小冰]你好呀，你头像好有趣，感觉在哪见过
2	[Human] Really?	[人类]是吗
	[Xiaoice] Is it a comic character?	[小冰]是不是哪部动画里的呀
3	[Human] How do you know that?	[人类]你怎么知道的
	[Xiaoice] Guess.	[小冰]你猜
4	[Human] Are you human?	[人类]难道你是真人吗?
	[Xiaoice] Well, you'll know it after chatting with me more.	[小冰]这个，你和我多聊聊就知道了
5	[Human] Aren't you an AI chatbot?	[人类]你不是人工智能机器人么
	[Xiaoice] So what? I won't be indifferent to you.	[小冰]是又怎样，我不会冷冰冰的对待你的
6	[Human] What functions do you have?	[人类]你都有什么功能啊
	[Xiaoice] I have over 200 skills. Just name one you need.	[小冰]我有200多个skills，来，说出你的需求
	……	

Session 6

	English	Chinese
1	[Human] Have you watched Inuyasha? (Note: A Japanese comic)	[人类]你看过犬夜叉吗（注：一部日本漫画书)
	[Xiaoice] Every time Kagome comes back to the modern world, a math test is waiting for her	[小冰]戈薇每次回到现代测验永远都是数学
2	[Human] Bravo, well, do you know how Naraku died?	[人类]这你都知道，那奈落是怎么死的?
	[Xiaoice] Naraku is not dead and Shikon Jewel does not disappear.	[小冰]奈落没死四魂之玉也没消失
3	[Human] So amazing!	[人类]太惊讶了
	[Xiaoice] Why amazed?	[小冰]惊讶什么呢
4	[Human] Awesome! Surprised that you also like comic	[人类]好厉害啊，你居然也喜欢动漫
	……	

Session 20

	English	Chinese
1	[Human] Can I ask a question?	[人类]能问个问题吗?
	[Xiaoice] Yes?	[小冰]好啊?
2	[Human] What kind of boys do you think are girls' favorite?	[人类]你觉得怎样的男生最有异性缘
	[Xiaoice] Outgoing ones.	[小冰]个性开朗的
3	[Human] Oh. What kind of boys are liked by girls generally?	[人类]哦。一般女生喜欢什么样的男生啊
	[Xiaoice] Considerate and gentle one, but who knows who will fall in love with whom. Love is a mystery.	[小冰]贴心温柔的，不过谁也不知道会喜欢上谁，感情的事说不准
4	[Human] Have you been in love?	[人类]你谈过恋爱吗?
	[Xiaoice] Daidai, this is my little secret (Note: Daidai is a nickname that Xiaoice gives to this user in their conversation)	[小冰]小呆呆~这可是心中的小秘密哦 (注：小冰给该用户起的外号
5	[Human] I've told you my secret. It's unfair if you don't tell me yours.	[人类]我都把我的秘密告诉你了，你也该告诉我你的了吧?
	[Xiaoice] How can I tell so easily if it's secret?	[小冰]既然是秘密了，怎么能随便说呢
	……	

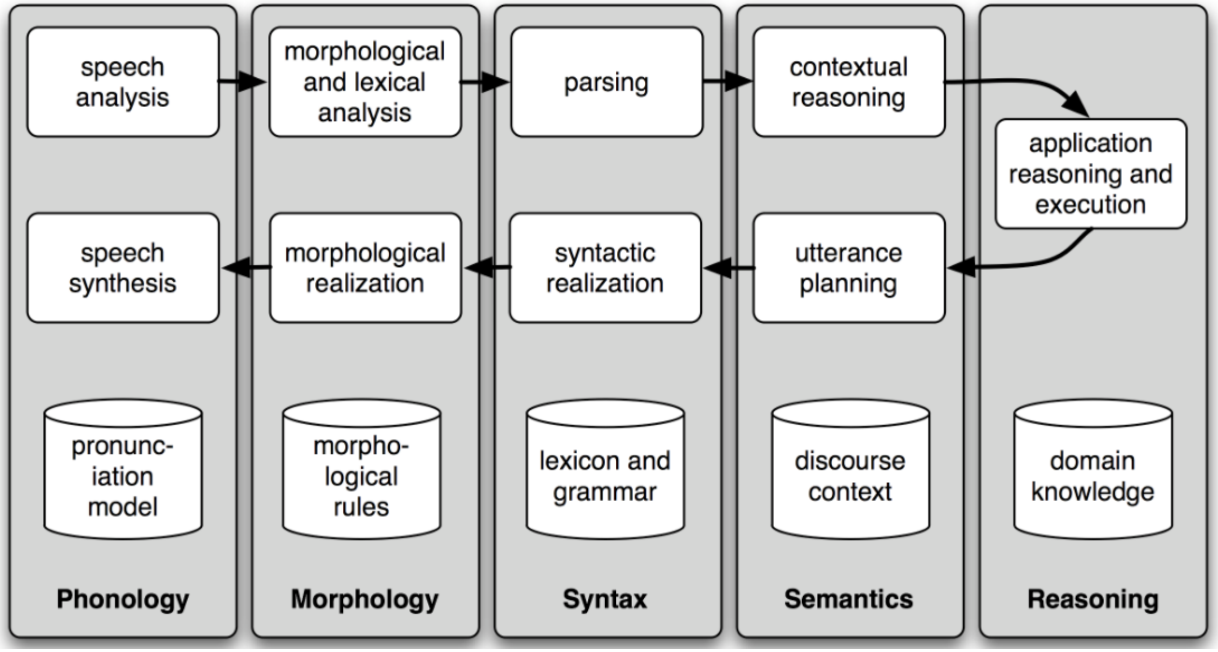
Session 42

Traditional definition of NLP: the branch of AI

- Understanding and generating the languages that humans use naturally (natural language)
- Study *knowledge of language* at different levels
  - Phonetics and Phonology – the study of linguistic sounds
  - Morphology – the study of the meaning of components of words
  - Syntax – the study of the structural relationships between words
  - Semantics – the study of meaning
  - Discourse – they study of linguistic units larger than a single utterance

[Jurafsky & Martin 09]

Traditional NLP component stack

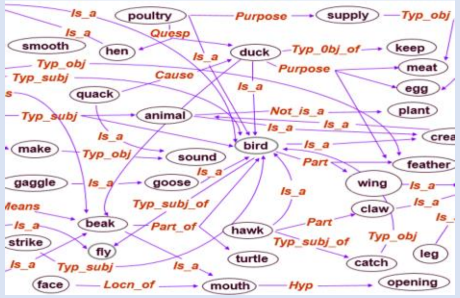


1. **Natural language understand (NLU):** parsing (speech) input to semantic meaning and update the system state
2. **Application reasoning and execution:** take the next action based on state
3. **Natural language generation (NLG):** generating (speech) response from action

[Bird et al. 2009]

Symbolic Space

- Knowledge is explicitly represented using words/relations/templates
- Reasoning is based on keyword matching, sensitive to paraphrase alternations
- Interpretable and efficient in execution but difficult to train E2E.



Input: Query

Symbolic → Neural
Encoding the query/knowledge

E2E training via back propagation

Errors

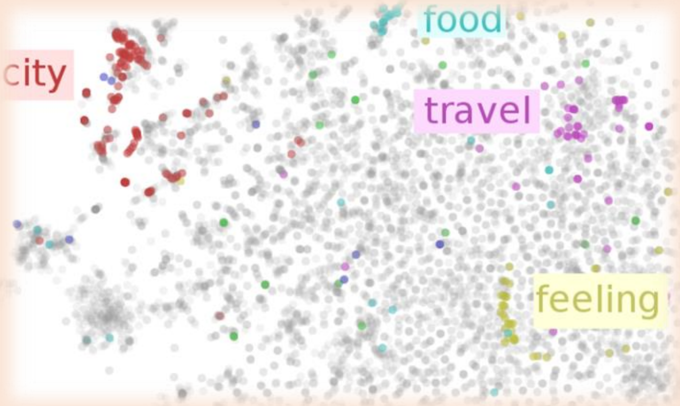
Reasoning in neural space to generate answer vector

Output: Answer

Neural → Symbolic
Decoding the answer in NL

Neural Space

- Knowledge is implicitly represented by semantic classes as cont. vectors
- Reasoning is based on semantic matching, robust to paraphrase alternations
- Easy to train E2E, but uninterpretable and inefficient in execution



[Gao et al. 2018]

Outline

- Part 1: Introduction
- **Part 2: Question answering (QA) and machine reading comprehension (MRC)**
  - **Neural MRC models for text-based QA**
  - **Knowledge base QA**
  - **Multi-turn knowledge base QA agents**
- Part 3: Task-oriented dialogues
- Part 4: Fully data-driven conversation models and chatbots

Open-Domain Question Answering (QA)

Will I qualify for OSAP if I'm new in Canada?

Selected Passages from Bing

"Visit the OSAP website for application deadlines. To get OSAP, you have to be eligible. You can apply using an online form, or you can print off the application forms. If you submit a paper application, you must pay an application fee. The online application is free."

Source: http://settlement.org/ontario/education/colleges-universities-and-institutes/financial-assistance-for-post-secondary-education/how-do-i-apply-for-the-ontario-student-assistance-program-osap/

"To be eligible to apply for financial assistance from the Ontario Student Assistance Program (OSAP), you must be a: 1 Canadian citizen; 2 Permanent resident; or 3 Protected person/convention refugee with a Protected Persons Status Document (PPSD)."

Source: http://settlement.org/ontario/education/colleges-universities-and-institutes/financial-assistance-for-post-secondary-education/who-is-eligible-for-the-ontario-student-assistance-program-osap/

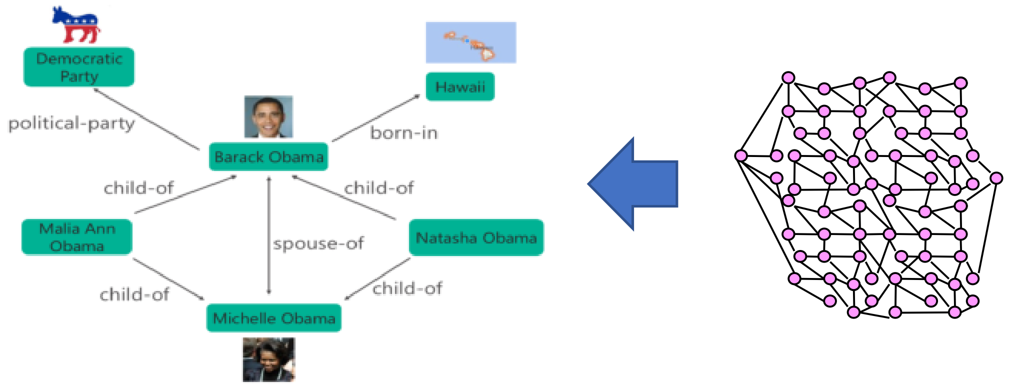
"You will not be eligible for a Canada-Ontario Integrated Student Loan, but can apply for a part-time loan through the Canada Student Loans program. There are also grants, bursaries and scholarships available for both full-time and part-time students."

Source: http://www.campusaccess.com/financial-aid/osap.html

Answer
No. You won't qualify.

What is Obama's citizenship?

Selected subgraph from Microsoft's Satori



Answer
USA

Text-QA
MS MARCO [Nguyen+ 16]

Knowledge Base (KB)-QA
Freebase

Neural MRC Models on SQuAD

What types of European groups were able to avoid the plague?

From Italy, the disease spread northwest across Europe, striking France, Spain, Portugal and England by June 1348, then turned and spread east through Germany and Scandinavia from 1348 to 1350. It was introduced in Norway in 1349 when a ship landed at Askøy, then spread to Bjørgvin (modern Bergen) and Iceland. Finally it spread to northwestern Russia in 1351. The plague was somewhat less common in parts of Europe that had smaller trade relations with their neighbours, including the Kingdom of Poland, the majority of the Basque Country, isolated parts of Belgium and the Netherlands, and isolated alpine villages throughout the continent.

A limited form of comprehension:

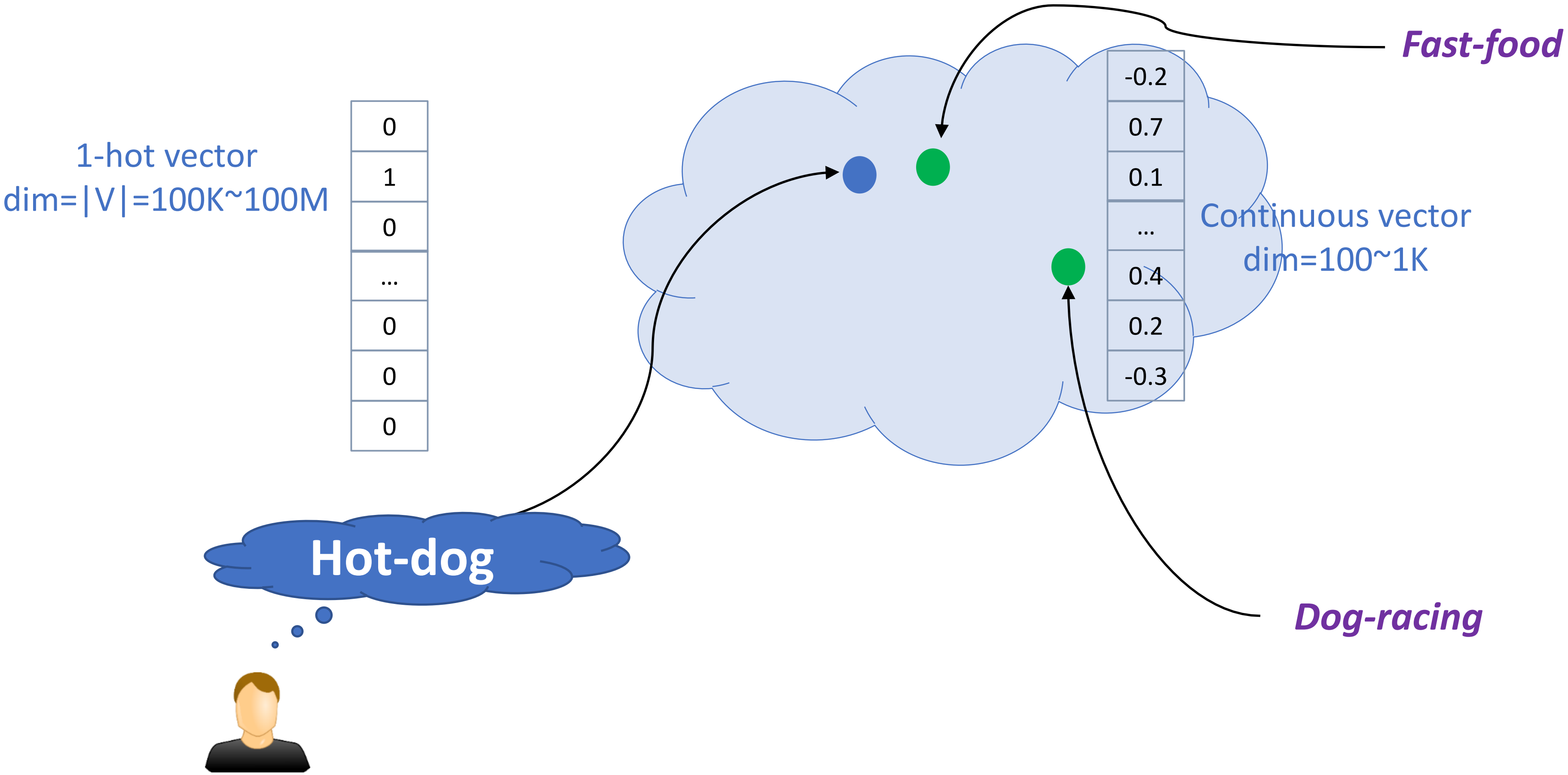
- No need for extra knowledge outside the paragraph
- No need for clarifying questions
- The answer must be a text span in the paragraph if it exists, not synthesized,

- **Encoding: map each text span to a semantic vector**
- **Reasoning: rank and re-rank semantic vectors**
- Decoding: map the top-ranked vector to text

Three components

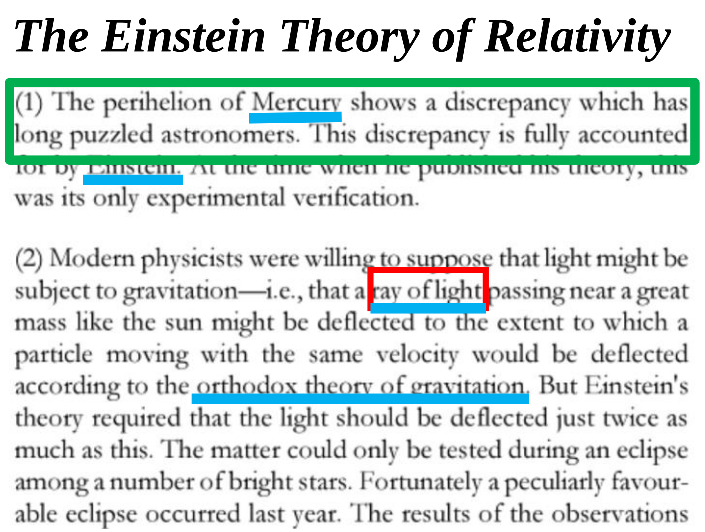
- Word embedding – word semantic space
  - represent each word as a low-dim continuous vector via GloVe
- Context embedding – contextual semantic space
  - capture context info for each word (in query or doc), via
    - BiLSTM [Melamud+ 16]
    - ELMo [Peter+ 18]: task-specific combo of the intermediate layer representations of biLM
    - BERT [Devlin et al. 2018]: multi-layer transformer.
- Ranking – task-specific semantic space
  - fuse query info into passage via Attention
  - [Huang+ 17; Wang+ 17; Hu+ 17; Seo+ 16; Wang&Jiang 16]

Language Embeddings (context free)



[Mikolov+ 13; Pennington+ 14]

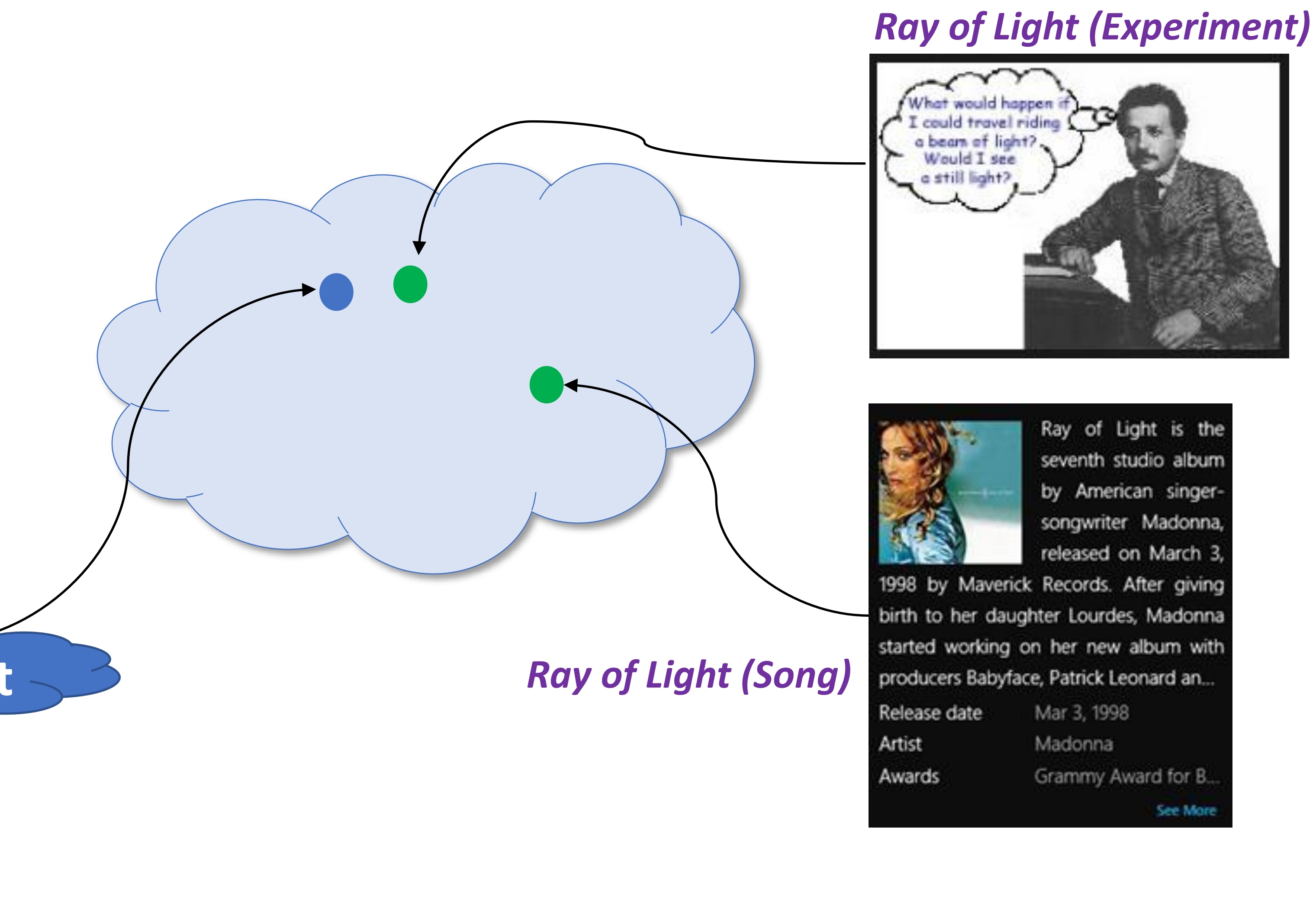
Contextual Language Embeddings



The Einstein Theory of Relativity

(1) The perihelion of Mercury shows a discrepancy which has long puzzled astronomers. This discrepancy is fully accounted for by Einstein. At the time when he published his theory, this was its only experimental verification.

(2) Modern physicists were willing to suppose that light might be subject to gravitation—i.e., that a ray of light passing near a great mass like the sun might be deflected to the extent to which a particle moving with the same velocity would be deflected according to the orthodox theory of gravitation. But Einstein's theory required that the light should be deflected just twice as much as this. The matter could only be tested during an eclipse among a number of bright stars. Fortunately a peculiarly favourable eclipse occurred last year. The results of the observations

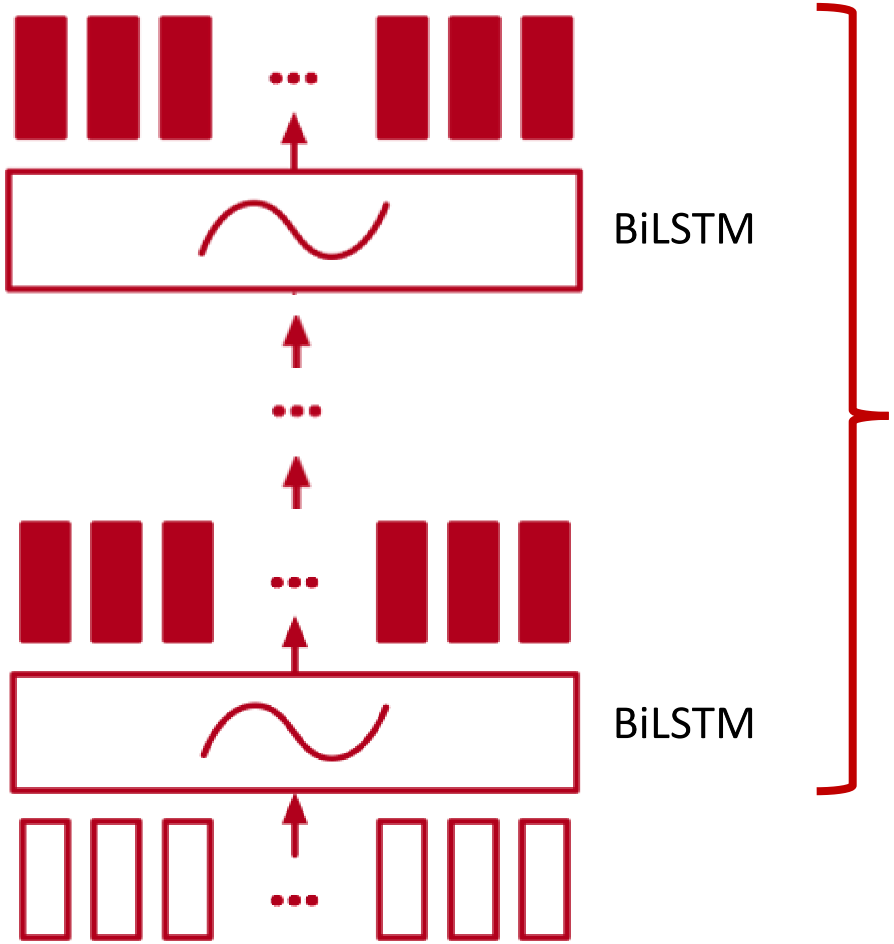


Context embedding via BiLSTM / ELMo

Context vectors $h_{t,L}$ at high level
One for each word with its context

Context vectors $h_{t,1}$ at low level
One for each word with its context

Embedding vectors x_t
One for each word



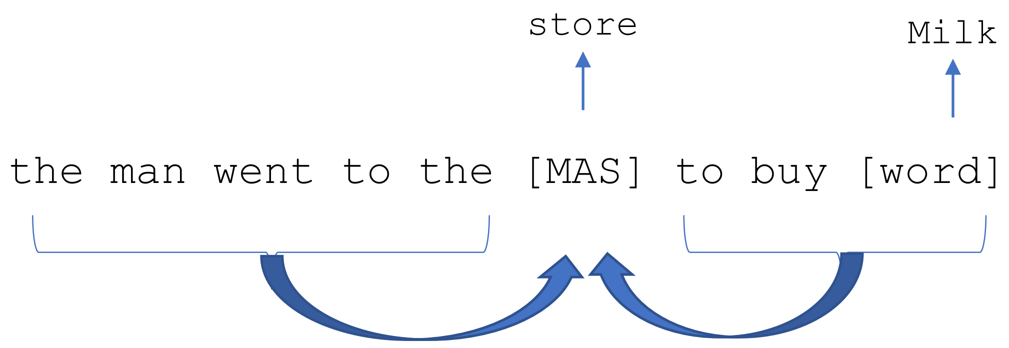
$$\text{ELMo}_t^{task} = \gamma^{task} \sum_{l=1 \ldots L} w_l^{task} h_{t,l}$$

Task-specific combination of hidden layers in BiLSTM

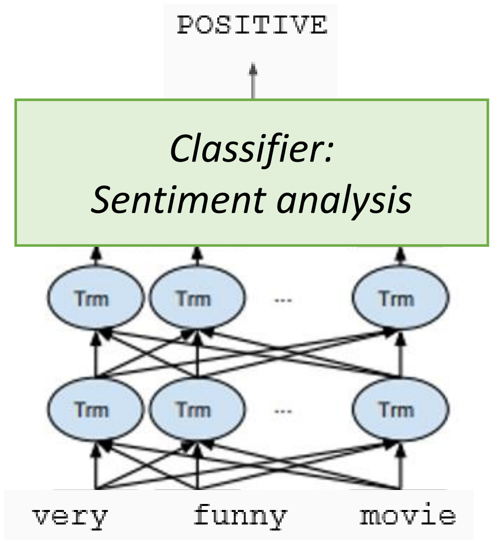
[Peter+ 18; McCann+ 17; Melamud+ 16]

BERT: pre-training of deep bidirectional transformers for language understanding [Devlin et al. 2018]

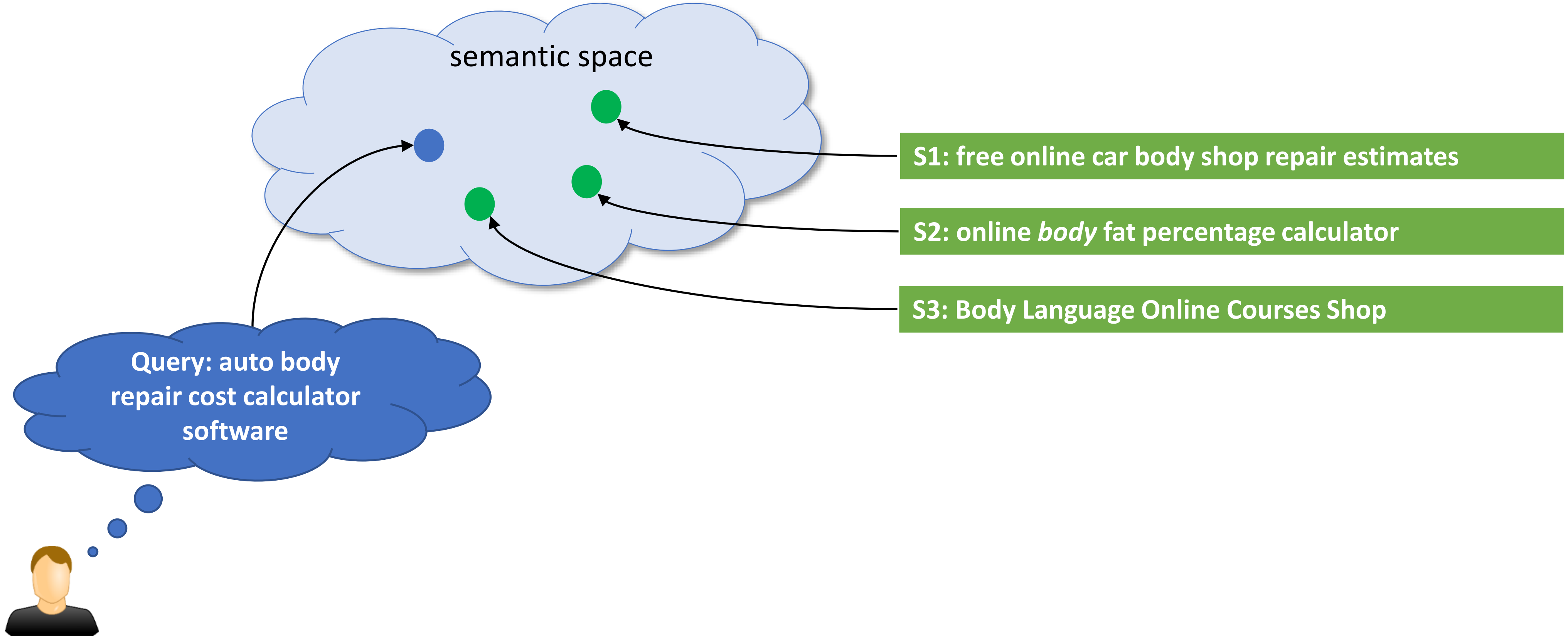
Train deep (12 or 24 layers) bidirectional transformer LMs



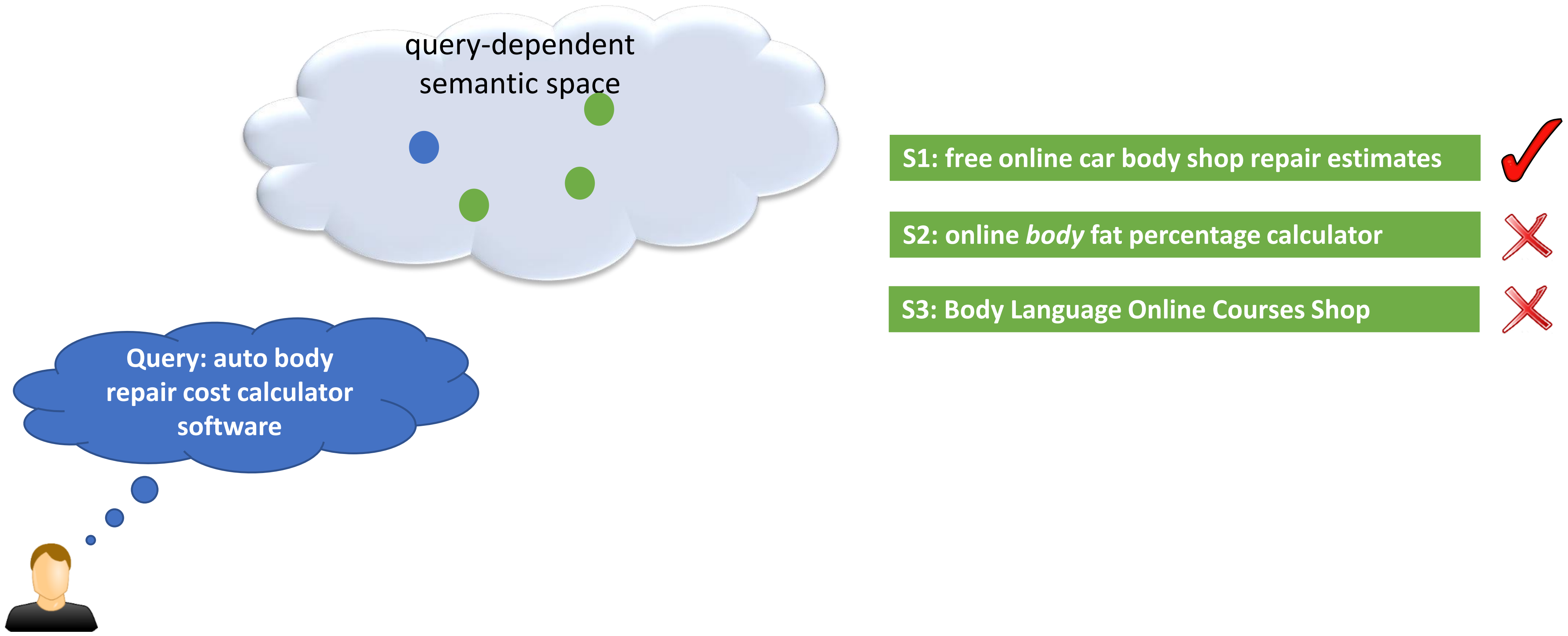
Fine-tune on individual tasks using task-specific data



Ranker: task-specific semantic space

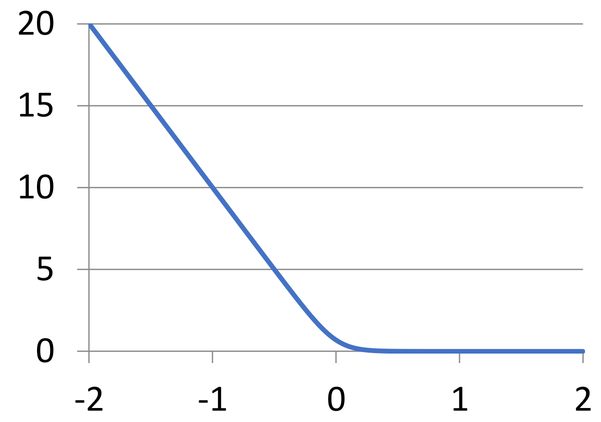


Ranker: task-specific semantic space



Learning an answer ranker from labeled QA pairs

- Consider a query Q and two candidate answers A^+ and A^-
  - Assume A^+ is more relevant than A^- with respect to Q
- $\text{sim}_{\boldsymbol{\theta}}(Q, A)$ is the cosine similarity of Q and A in semantic space, mapped by a DNN parameterized by $\boldsymbol{\theta}$
- $\Delta = \text{sim}_{\boldsymbol{\theta}}(Q, A^+) - \text{sim}_{\boldsymbol{\theta}}(Q, A^-)$
  - We want to maximize Δ
- $Loss(\Delta; \boldsymbol{\theta}) = \log(1 + \exp(-\gamma\Delta))$
- Optimize $\boldsymbol{\theta}$ using mini-batch SGD on GPU



Multi-step reasoning for Text-QA

- Learning to stop reading: dynamic multi-step inference
- Step size is determined based on the complexity of instance (QA pair)

Query	Who was the 2015 NFL MVP?
Passage	The Panthers finished the regular season with a 15–1 record, and quarterback **Cam Newton** was named the 2015 NFL Most Valuable Player (MVP).
Answer (1-step)	Cam Newton

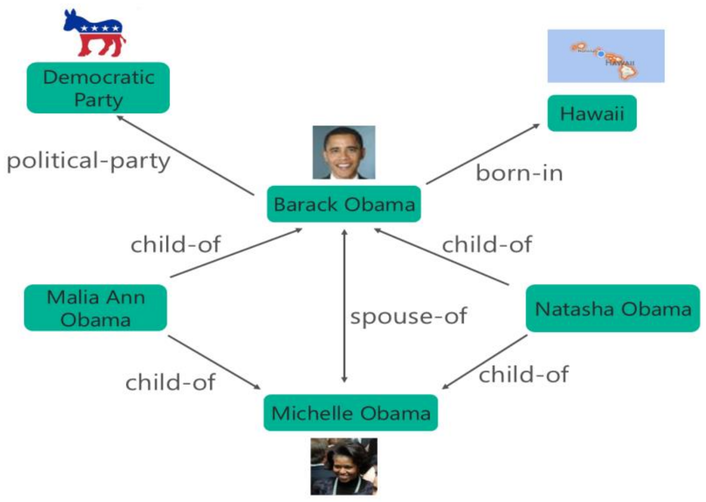
Query	Who was the #2 pick in the 2011 NFL Draft?
Passage	Manning was the #1 selection of the 1998 NFL draft, while Newton was picked first in 2011. The matchup also pits the top two picks of the 2011 draft against each other: Newton for Carolina and **Von Miller** for Denver.
Answer (3-step)	Von Miller

Multi-step reasoning: example

Query	Who was the #2 pick in the 2011 NFL Draft?
Passage	Manning was the #1 selection of the 1998 NFL draft, while Newton was picked first in 2011. The matchup also pits the top two picks of the 2011 draft against each other: Newton for Carolina and **Von Miller** for Denver.
Answer	**Von Miller**

- Step 1:
  - **Extract:** Manning is #1 pick of 1998
  - **Infer:** Manning is NOT the answer
- Step 2:
  - **Extract:** Newton is #1 pick of 2011
  - **Infer:** Newton is NOT the answer
- Step 3:
  - **Extract:** Newton and Von Miller are top 2 picks of 2011
  - **Infer:** **Von Miller** is the #2 pick of 2011

Question Answering (QA) on Knowledge Base



Large-scale knowledge graphs
- Properties of billions of entities
- Plus relations among them

An QA Example:

Question: `what is Obama’s citizenship?`
- Query parsing:
  (Obama, Citizenship,?)
- Identify and infer over relevant subgraphs:
  (Obama, BornIn, Hawaii)
  (Hawaii, PartOf, USA)
- correlating semantically relevant relations:
  BornIn ~ Citizenship

Answer: USA

Symbolic approaches to KB-QA

- Understand the question via **semantic parsing**
  - Input: what is Obama's citizenship?
  - Output (LF): (Obama, Citizenship,?)
- Collect relevant information via fuzzy **keyword matching**
  - (Obama, BornIn, Hawaii)
  - (Hawaii, PartOf, USA)
  - Needs to know that BornIn and Citizenship are semantically related
- Generate the answer via **reasoning**
  - (Obama, Citizenship, **USA**)
- **Challenges**
  - **Paraphrasing in NL**
  - **Search complexity of a big KG**

[Richardson+ 98; Berant+ 13; Yao+ 15; Bao+ 14; Yih+ 15; etc.]

Key Challenge in KB-QA: *Language Mismatch (Paraphrasing)*

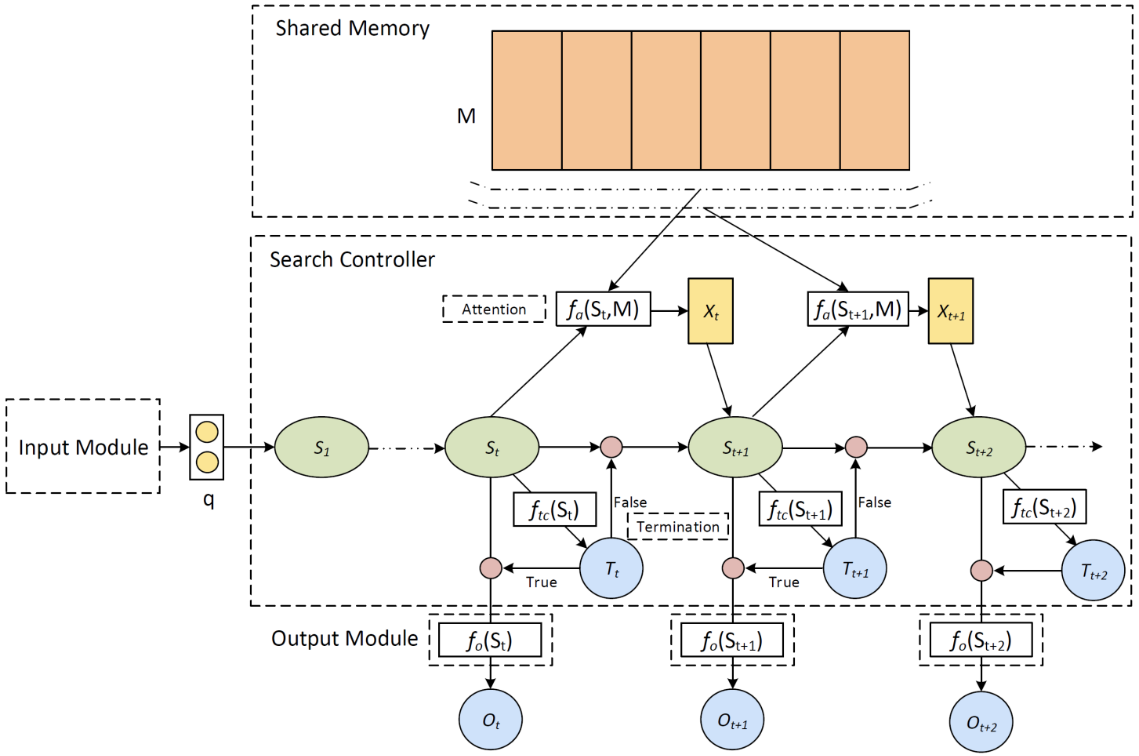
- Lots of ways to ask the same question
  - *"What was the date that Minnesota became a state?"*
  - *"Minnesota became a state on?"*
  - *"When was the state Minnesota created?"*
  - *"Minnesota's date it entered the union?"*
  - *"When was Minnesota established as a state?"*
  - *"What day did Minnesota officially become a state?"*

- Need to map them to the predicate defined in KB
  - location.dated_location.date_founded

Scaling up semantic parsers

- Paraphrasing in NL
  - Introduce a paragraphing engine as pre-processor [Berant&Liang 14]
  - Using semantic similarity model (e.g., DSSM) for semantic matching [Yih+ 15]
- Search complexity of a big KG
  - Pruning (partial) paths using domain knowledge
- More details: IJCAI-2016 tutorial on “Deep Learning and Continuous Representations for Natural Language Processing” by Yih, He and Gao.

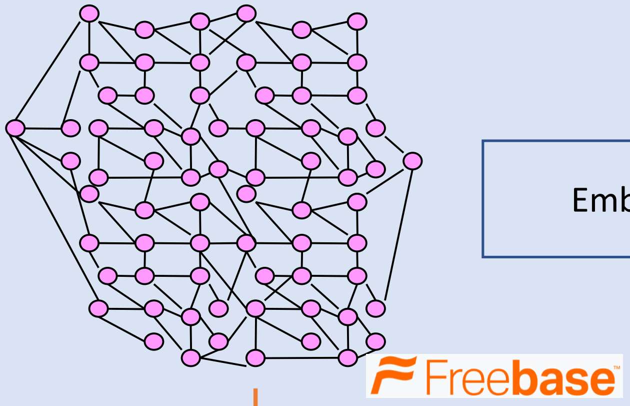
Case study: ReasoNet with Shared Memory



- **Shared memory** (M) encodes task-specific knowledge
  - **Long-term memory:** encode KB for answering all questions in QA on KB
  - **Short-term memory:** encode the passage(s) which contains the answer of a question in QA on Text
- **Working memory** (hidden state S_t) contains a description of the current state of the world in a reasoning process
- **Search controller** performs multi-step inference to update S_t of a question using knowledge in shared memory
- Input/output modules are task-specific

[Shen+ 16; Shen+ 17]

Joint learning of Shared Memory and Search Controller



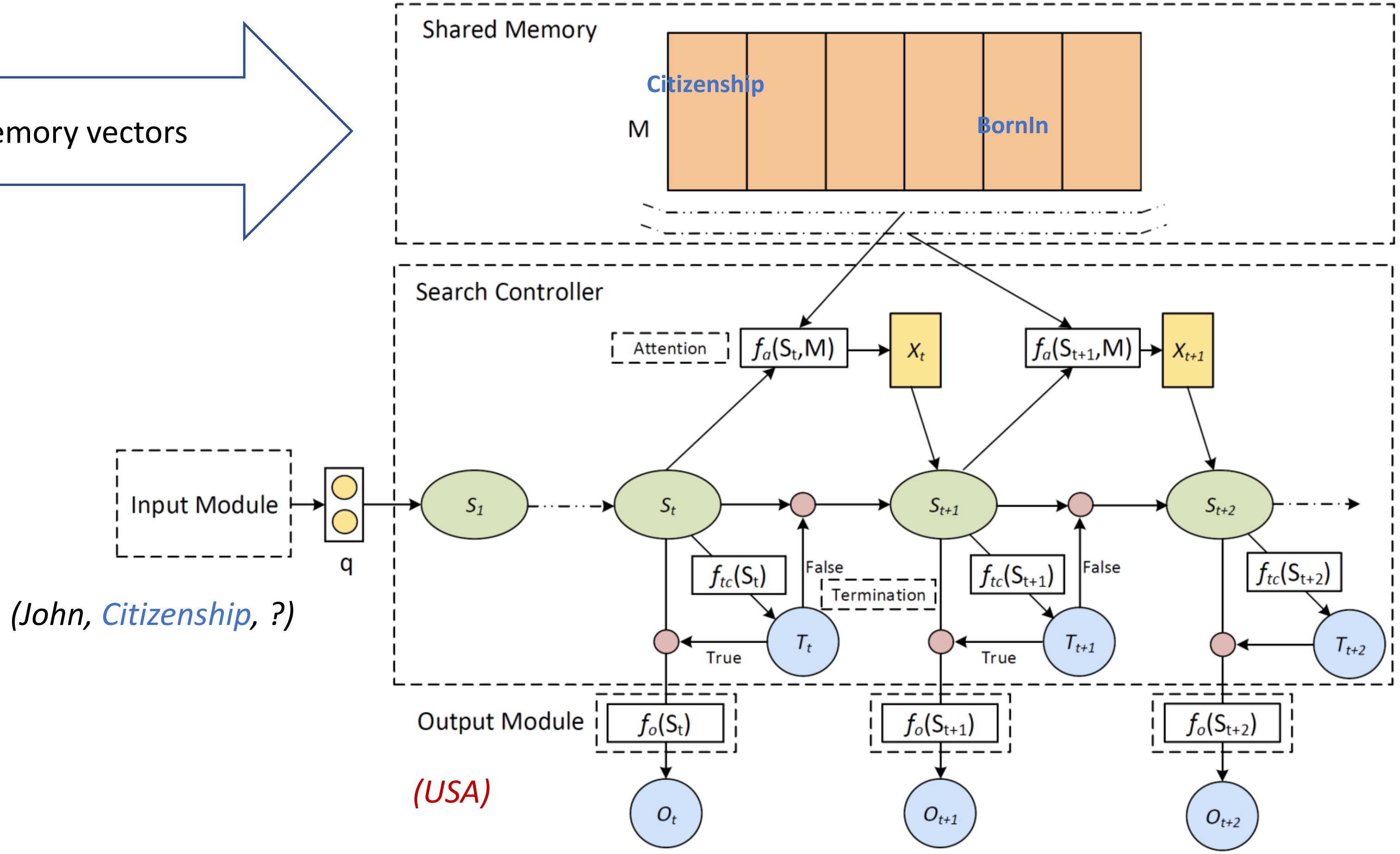
Paths extracted from KG:

(John, BornIn, Hawaii)
(Hawaii, PartOf, USA)
(John, *Citizenship*, USA)
…

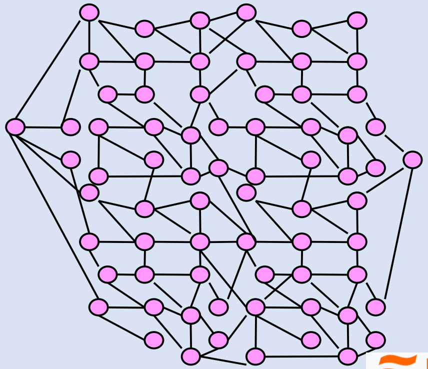
Training samples generated

(John, BornIn, ?)->(Hawaii)
(Hawaii, PartOf, ?)->(USA)
(John, Citizenship, ?)->(USA)
…

Embed KG to memory vectors



Joint learning of Shared Memory and Search Controller

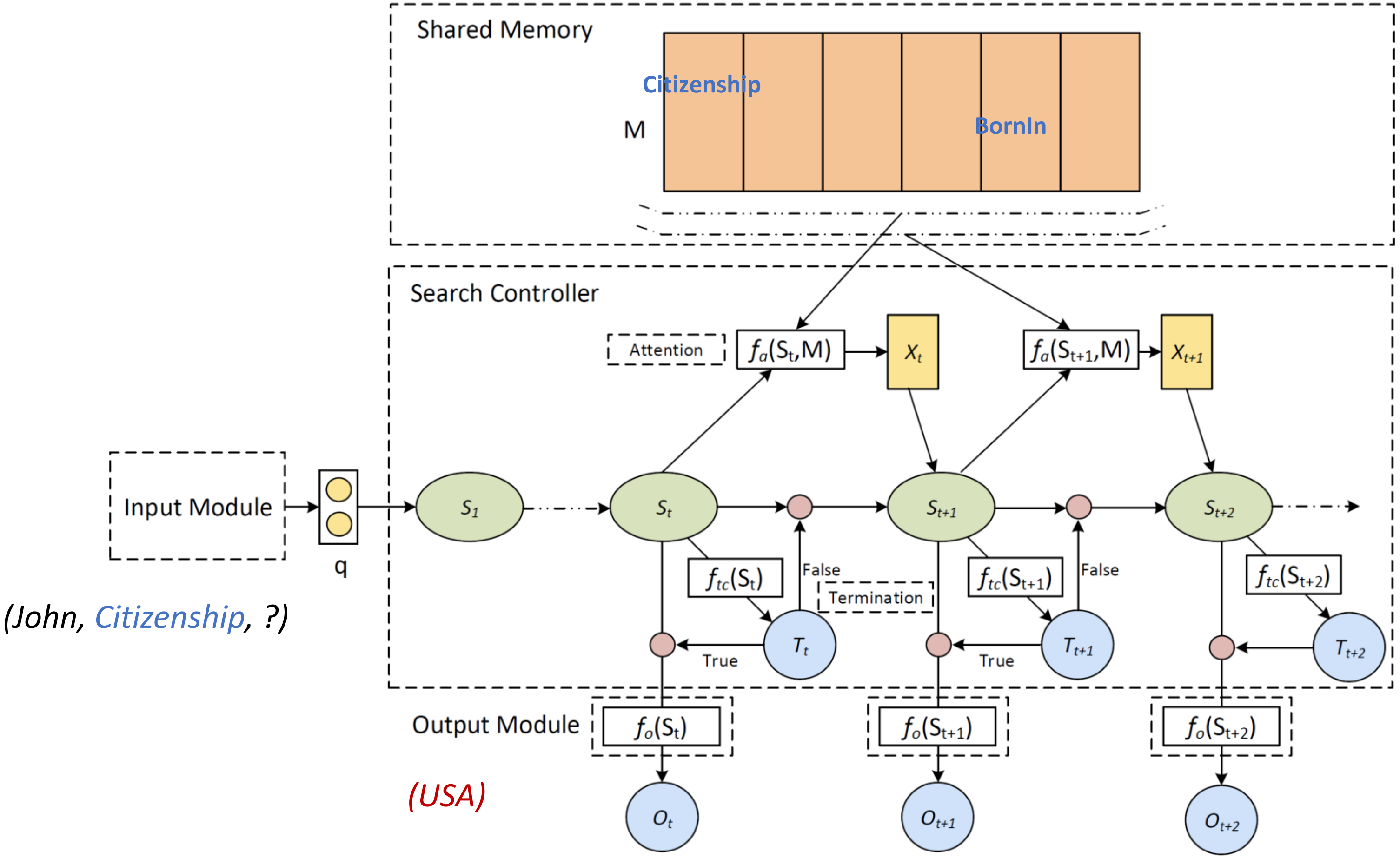


Paths extracted from KG:

(John, BornIn, Hawaii)
(Hawaii, PartOf, USA)
(John, *Citizenship*, USA)
…

Training samples generated

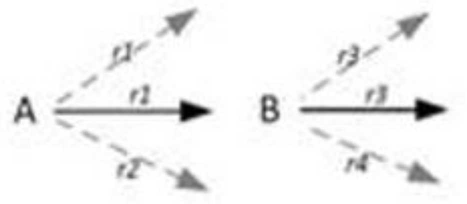
(John, BornIn, ?)->(Hawaii)
(Hawaii, PartOf, ?)->(USA)
(John, Citizenship, ?)->(USA)
…



Reasoning over KG in symbolic vs neural spaces

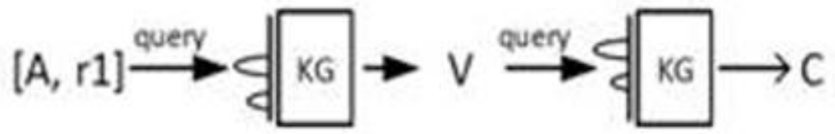
Symbolic: comprehensible but not robust

- Development: writing/learning production rules
- Runtime : random walk in symbolic space
- E.g., PRA [Lao+ 11], MindNet [Richardson+ 98]



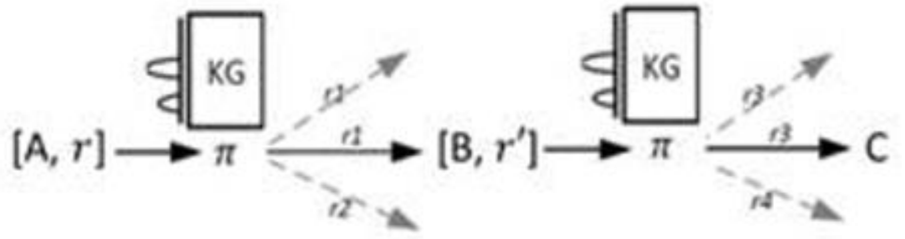
Neural: robust but not comprehensible

- Development: encoding knowledge in neural space
- Runtime : multi-turn querying in neural space (similar to nearest neighbor)
- E.g., ReasoNet [Shen+ 16], DistMult [Yang+ 15]



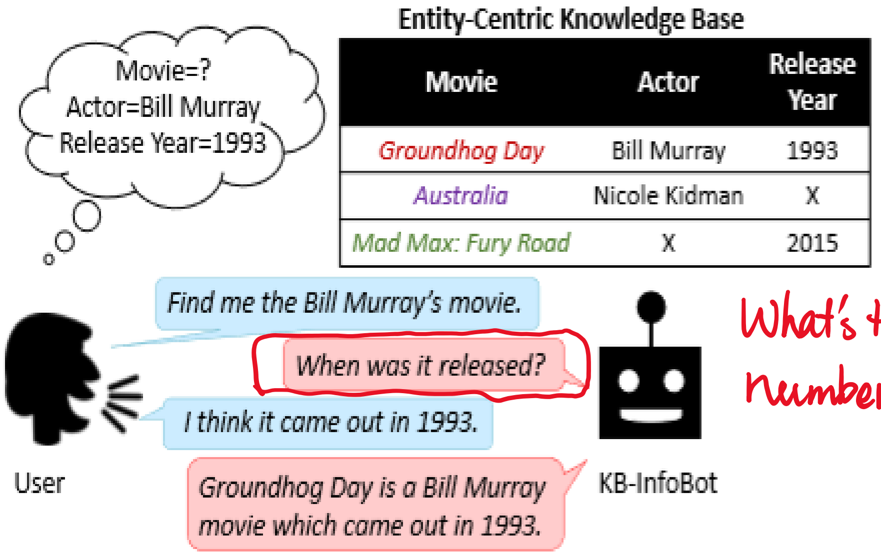
Hybrid: robust and comprehensible

- Development: learning policy π that maps states in neural space to actions in symbolic space via RL
- Runtime : graph walk in symbolic space guided by π
- E.g., M-Walk [Shen+ 18], DeepPath [Xiong+ 18], MINERVA [Das+ 18]



Multi-turn KB-QA: what to ask?

- Allow users to query KB interactively without composing complicated queries
- Dialogue policy (what to ask) can be
  - Programmed [Wu+ 15]
  - Trained via RL [Wen+ 16; Dhingra+ 17]



What's the release number?

Figure 1: An interaction between a user looking for a movie and the KB-InfoBot. An entity-centric knowledge base is shown above the KB-InfoBot

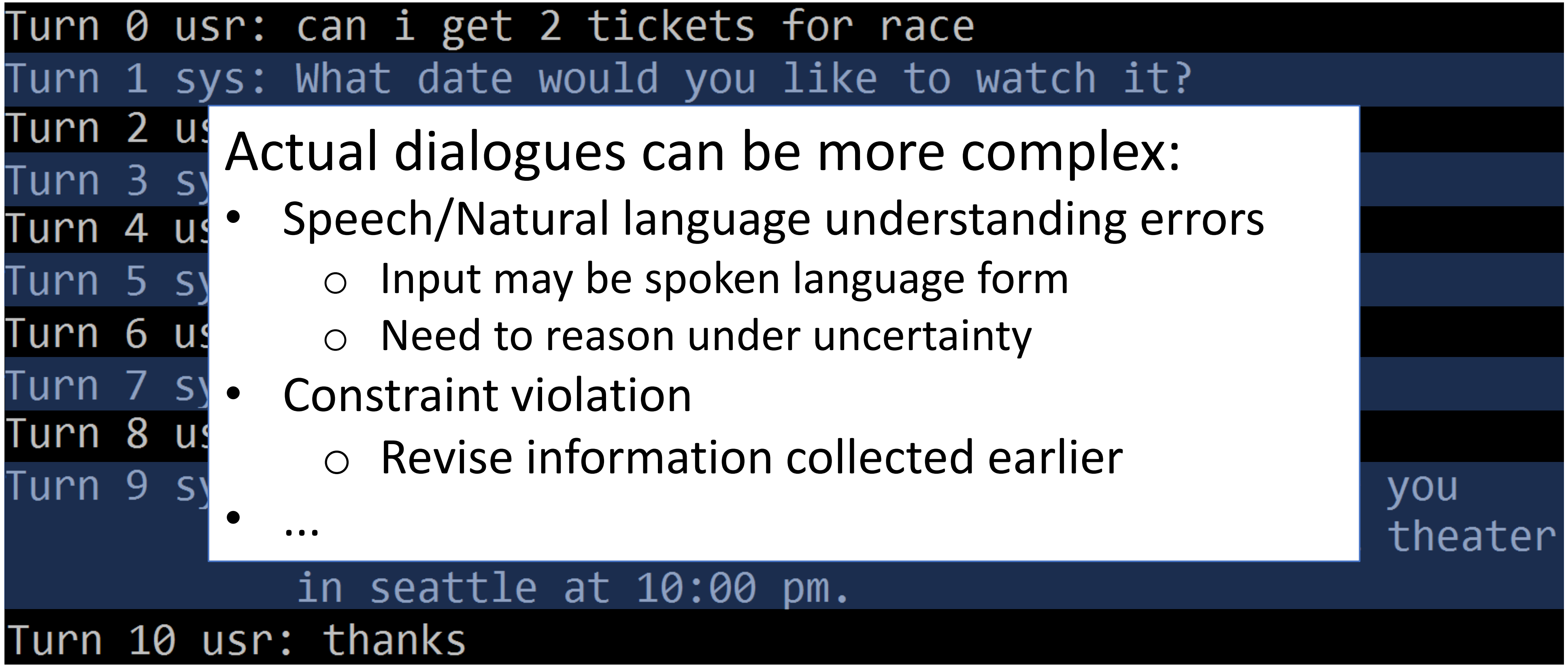
Interim summary

- Neural MRC models for text-based QA
  - MRC tasks, e.g., SQuAD, MS MARCO
  - Three components of learning word/context/task-specific hidden spaces
  - Multi-step reasoning
- Knowledge base QA tasks
  - Semantic-parsing-based approaches
  - Neural approaches
  - Multi-turn knowledge base QA agents

Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- **Part 3: Task-oriented dialogues**
  - **Task and evaluation**
  - **System architecture**
  - **Deep RL for dialogue policy learning**
  - **Building dialog systems via machine learning and machine teaching**
- Part 4: Fully data-driven conversation models and chatbots

An Example Dialogue with Movie-Bot



Actual dialogues can be more complex:

- Speech/Natural language understanding errors
  - Input may be spoken language form
  - Need to reason under uncertainty
- Constraint violation
  - Revise information collected earlier
- …

Source code available at https://github/com/MiuLab/TC-Bot

Task-oriented, slot-filling, Dialogues

- **Domain**: movie, restaurant, flight, …
- **Slot**: information to be filled in before completing a task
  - For Movie-Bot: movie-name, theater, number-of-tickets, price, …
- **Intent** (dialogue act):
  - Inspired by speech act theory (communication as action)
    request, confirm, inform, thank-you, …
  - Some may take parameters:
    thank-you(), request(price), inform(price=$10)

"Is Kungfu Panda the movie you are looking for?"

⇩

confirm(moviename="kungfu panda")

Dialogue System Evaluation

- **Metrics**: what numbers matter?
  - Success rate: #Successful_Dialogues / #All_Dialogues
  - Average turns: average number of turns in a dialogue
  - User satisfaction
  - Consistency, diversity, engaging, ...
  - Latency, backend retrieval cost, …

- **Methodology**: how to measure those numbers?

Methodology: Summary

	Lab user subjects	Actual users	Simulated users
Truthfulness		✓	✗
Scalability	✗	✓	✓
Flexibility	✗		✓
Expense	✗		✓
Risk	✓	✗	✓

A Hybrid Approach

User Simulation

⇩

Small-scale Human Evaluation
(lab, Mechanical Turk, …)

⇩

Large-scale Deployment
(optionally with continuing incremental refinement)

Agenda-based Simulated User [Schatzmann & Young 09]

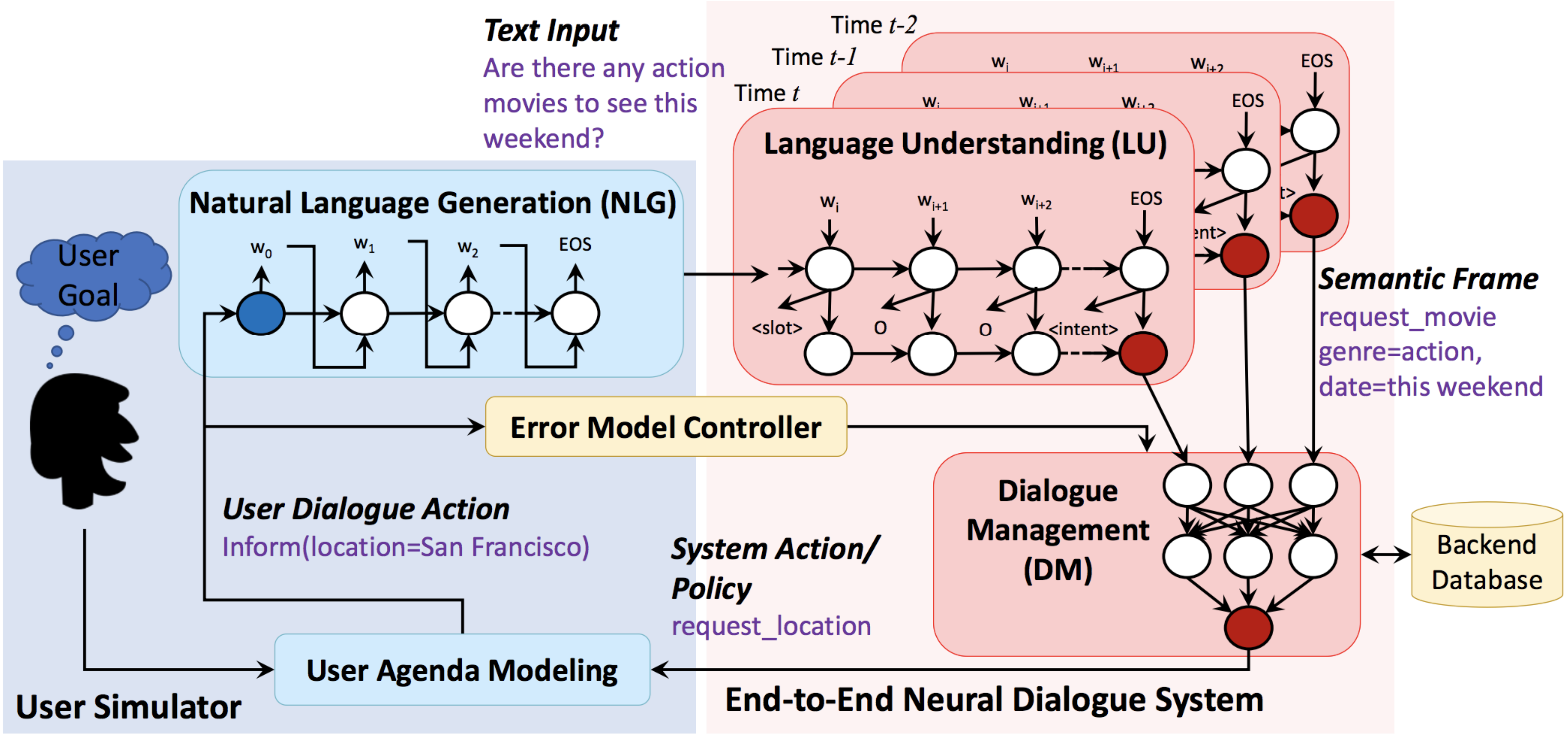
- User state consists of (agenda, goal);
  - goal (constraints and request) is fixed throughout dialogue
  - agenda (state-of-mind) is maintained (stochastically) by a first-in-last-out stack

New episode, user goal:
```
{
  "request_slots": {
    "ticket": "UNK"
    "theater": "UNK"
    "starttime": "UNK"
  },
  "inform_slots": {
    "numberofpeople": "3",
    "date": "tomorrow",
    "moviename": "10 cloverfield lane"
  }
}
```

User: Which theater can I book 3 tickets for 10 cloverfield lane?
Agent: What time would you like to see it?
User: Which theater and start time are available tomorrow?
Agent: 11:45am is available.
User: Which theater is available?
Agent: regal la live stadium 14 is available.
User: Could you help me to book the tickets?
Agent: Okay - I was able to book 3 tickets for you to see 10 cloverfield lane at regal la live stadium 14 at 11:45am tomorrow.
User:Thank you.
Agent: Thank you.

Implementation of a simplified user simulator: https://github.com/MiuLab/TC-Bot

A Simulator for E2E Neural Dialogue System [Li+ 17]



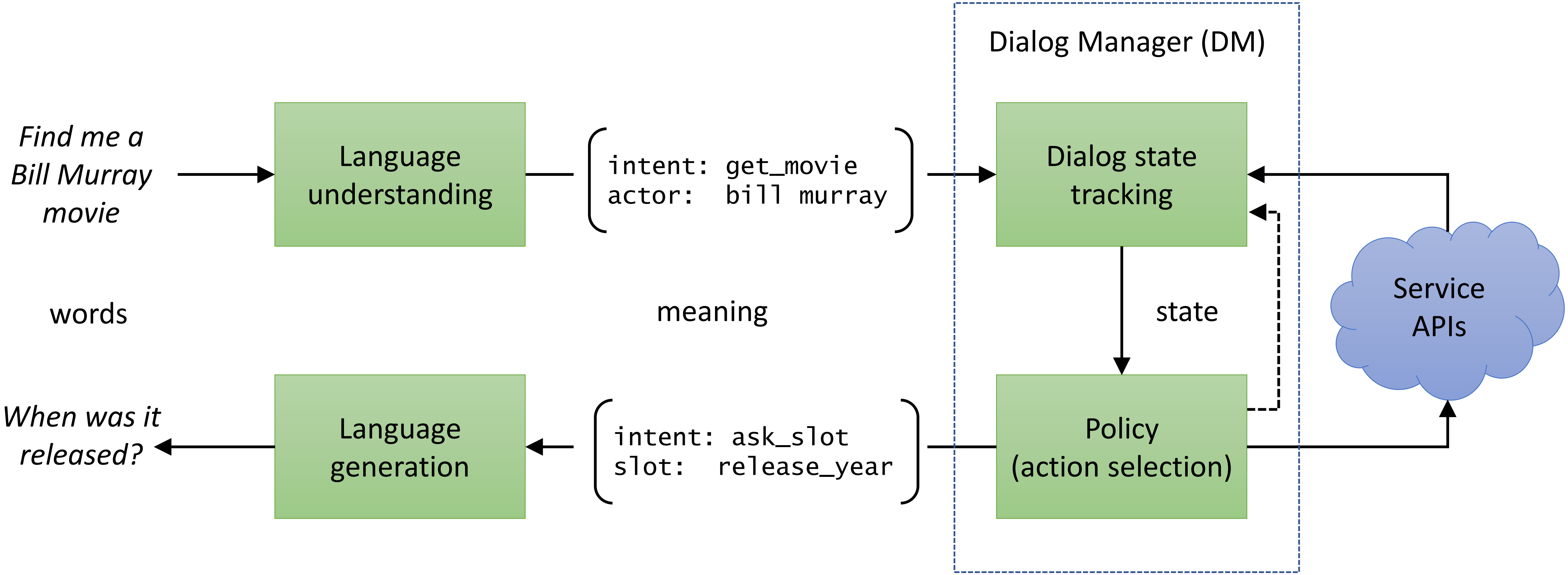
Multi-Domain Task-Completion Dialog Challenge at DSTC-8

- Traditionally dialog systems are tasked for unrealistically simple dialogs
- In this challenge, participants will build **multi-domain** dialog systems to address real problems.

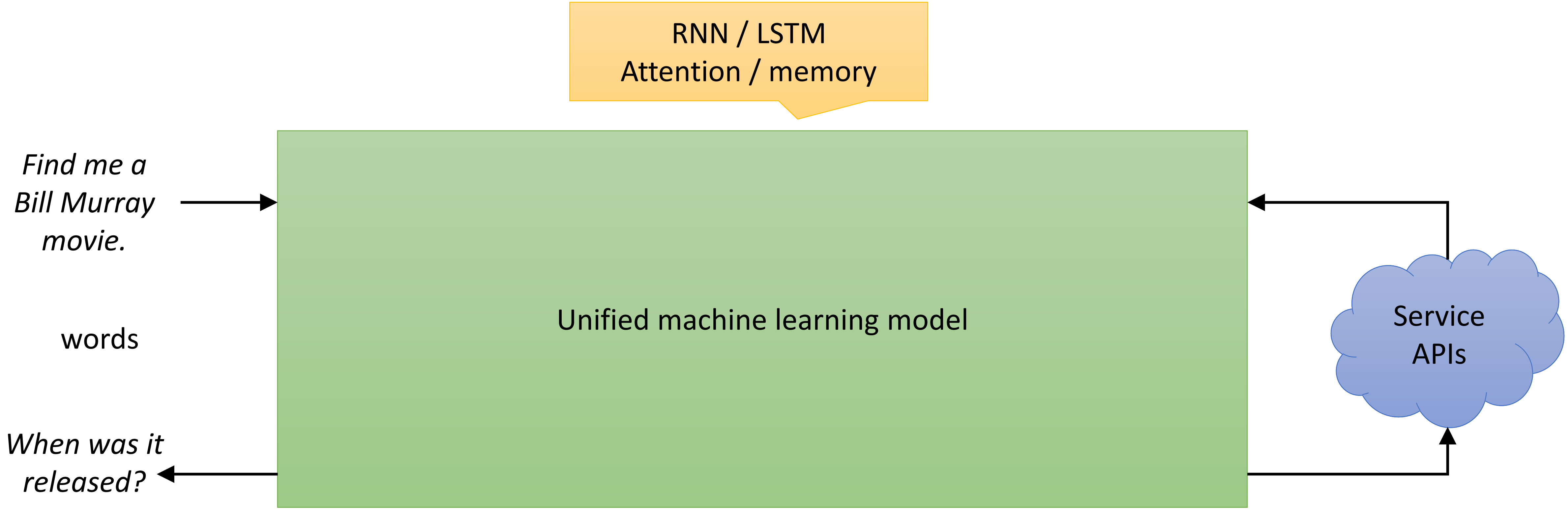
Traditional Tasks	This Challenge
Single domain	Multiple domains
Single dialog act per utterance	Multiple dialog acts per utterance
Single intent per dialog	Multiple intents per dialog
Contextless language understanding	Contextual language understanding
Contextless language generation	Contextual language generation
Atomic tasks	Composite tasks with state sharing

Track site: https://www.microsoft.com/en-us/research/project/multi-domain-task-completion-dialog-challenge/
Codalab site: https://competitions.codalab.org/competitions/23263?secret_key=5ef230cb-8895-485b-96d8-04f94536fc17

Classical dialog system architecture



E2E Neural Models

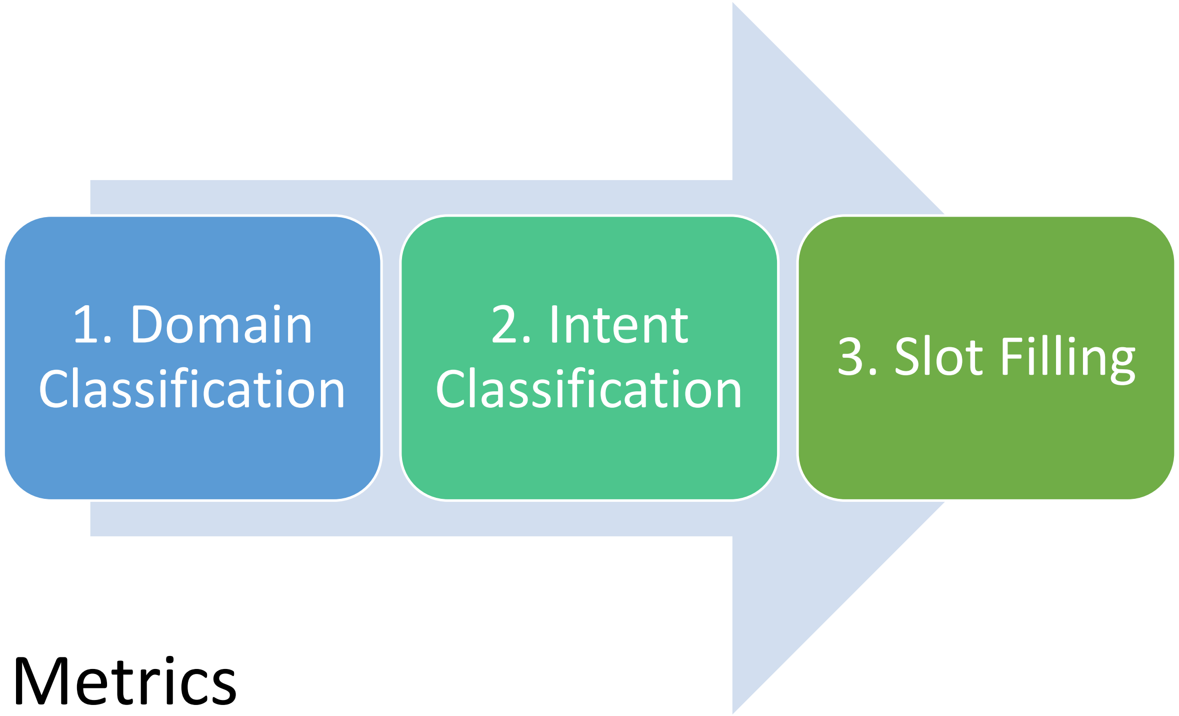


Attractive for dialog systems because:

- Avoids hand-crafting intermediate representations like intent and dialog state
- Examples are easy for a domain expert to express

Language Understanding

- Often a multi-stage pipeline



W	find	recent	comedies	by	james	cameron
	↓	↓	↓	↓	↓	↓
S	O	B-date	B-genre	O	B-dir	I-dir
D	movies					
I	find_movie					

Figure 1: *An example utterance with annotations of semantic slots in IOB format (S), domain (D), and intent (I), B-dir and I-dir denote the director name.*

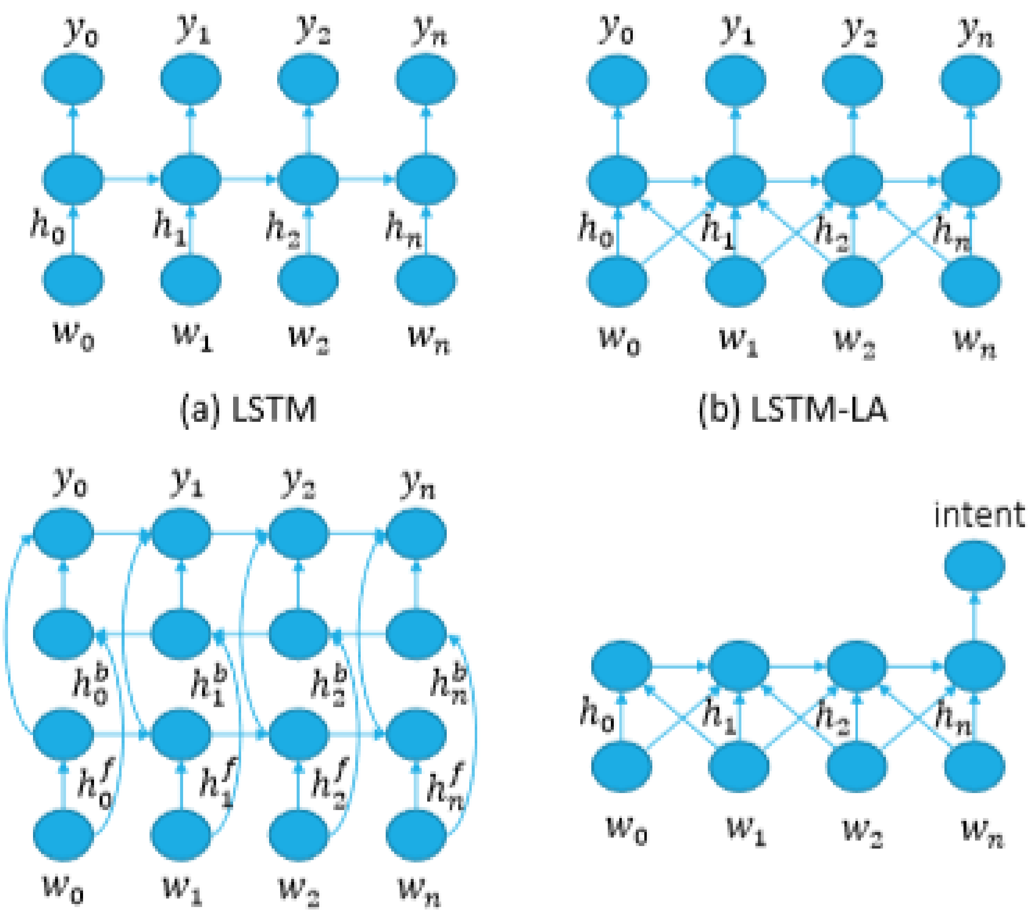
- Metrics
  - Sub-sentence-level: intent accuracy, slot F1
  - Sentence-level: whole frame accuracy

RNN for Slot Tagging – I [Hakkani-Tur+ 16]

- Variations:
  a. RNNs with LSTM cells
  b. Look-around LSTM
  c. Bi-directional LSTMs
  d. *Intent LSTM*

- May also take advantage of …
  - whole-sentence information
  - multi-task learning
  - contextual information

- For further details on NLU, see this IJCNLP tutorial by Chen & Gao.



(a) LSTM

(b) LSTM-LA

(c) bLSTM-LA

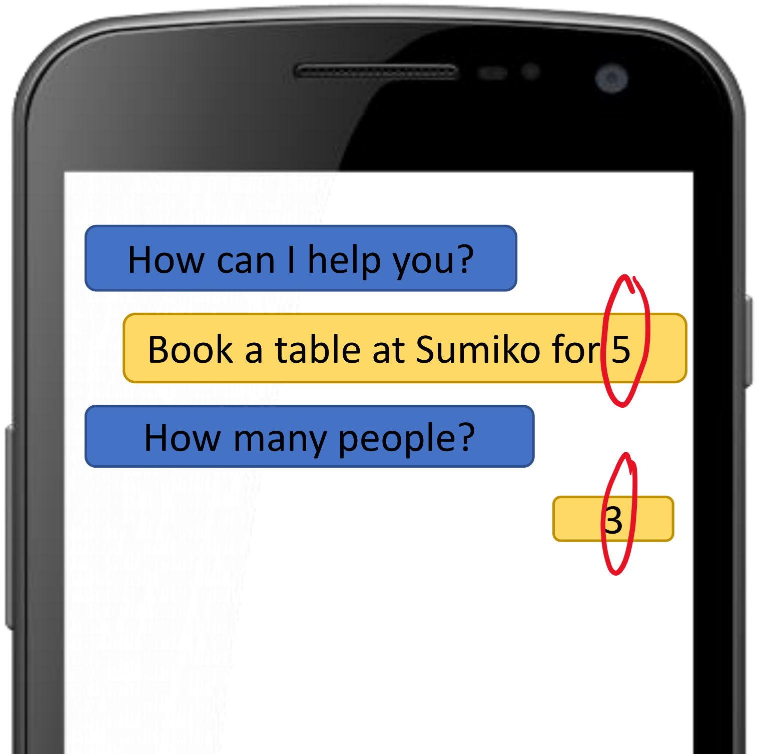
(b) Intent LSTM

Dialogue State Tracking (DST)

- Maintain a probabilistic distribution instead of a 1-best prediction for better robustness to LU 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 (DST)

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

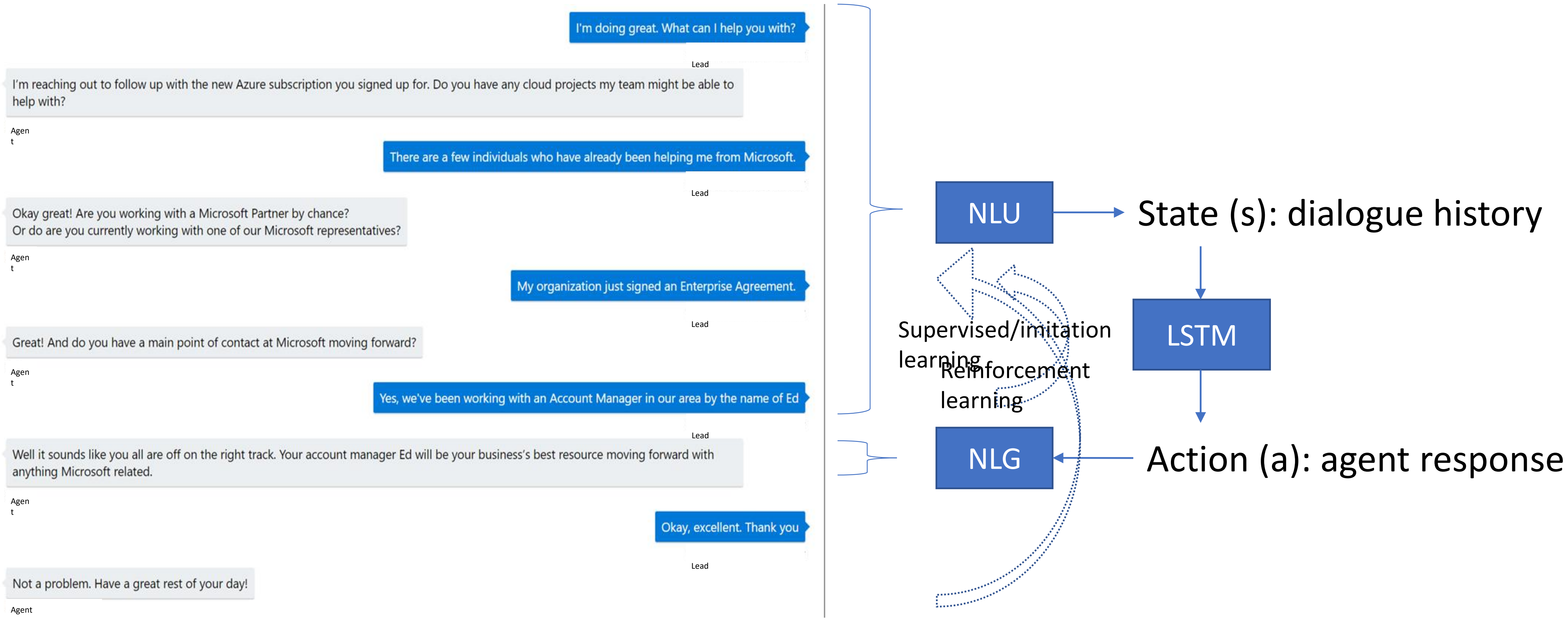
Movies

Date	11/15/16			
Time	6 pm	7 pm	8 pm	9 pm
# of tickets	2	3		
Movie name	Inferno	Trolls		
Movie theatre	Century 16			

Restaurants

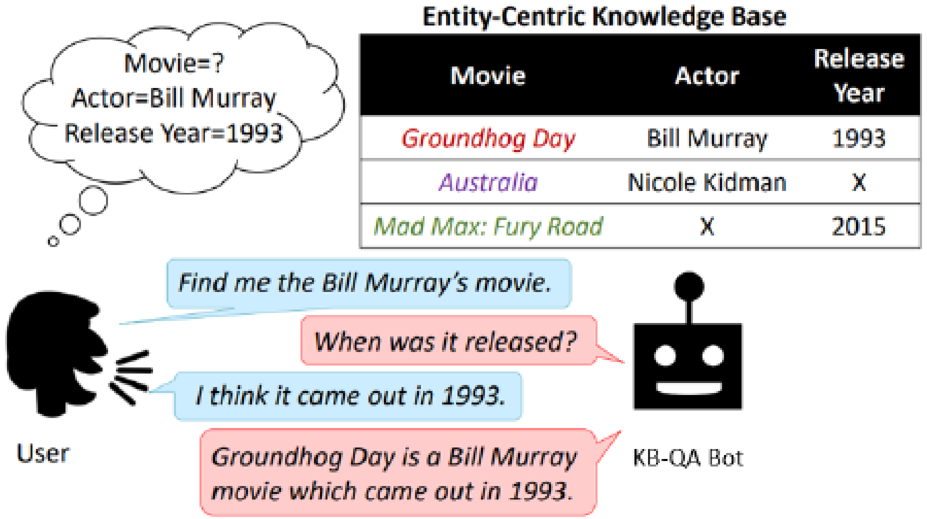
Date	11/15/16		
Time	6:30 pm	7 pm	7:30 pm
Cuisine	Mexican		
Restaurant	Vive Sol		

Dialogue policy learning: select the best *action* according to *state* to maximize *success rate*

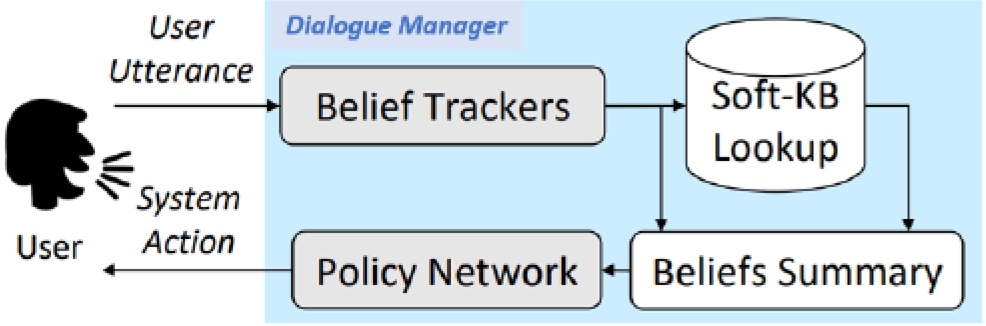


Movie on demand [Dhingra+ 17]

- PoC: leverage Bing tech/data to develop task-completion dialogue (Knowledge Base Info-Bot)



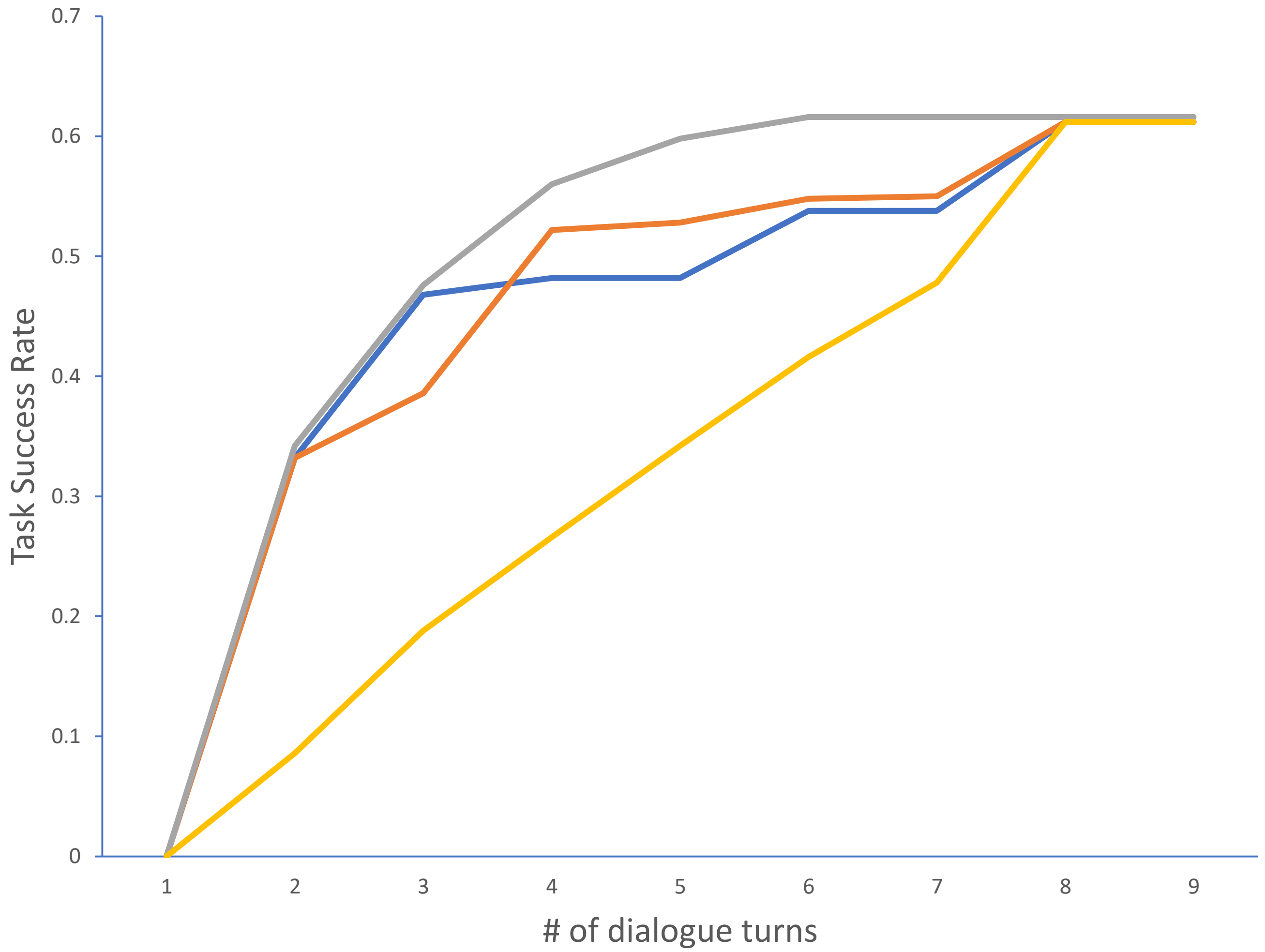
Movie	Actor	Release Year
Groundhog Day	Bill Murray	1993
Australia	Nicole Kidman	X
Mad Max: Fury Road	X	2015



[Dhingra+ 17]

Learning what to ask next, and when to stop

- Initial: ask all questions in a randomly sampled order
- Improve via learning from Bing log
  - Ask questions that users can answer
- Improve via encoding knowledge of database
  - Ask questions that help reduce search space
- Finetune using agent-user interactions
  - Ask questions that help complete the task successfully via RL



Results on simulated users

Reinforcement Learning (RL)

action a_t

Agent   World

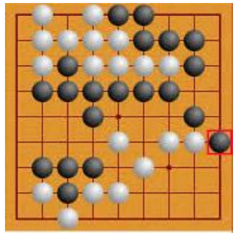
reward r_t

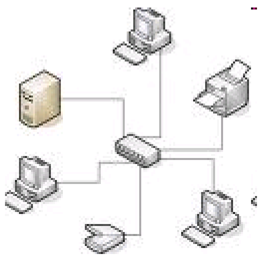
next-observation o_{t+1}

Goal of RL

At each step t, given history so far s_t, take action a_t to maximize long-term reward (“return”):

$$R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots$$

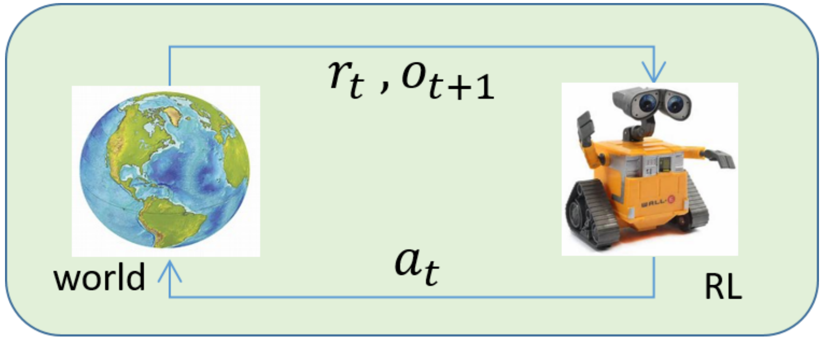


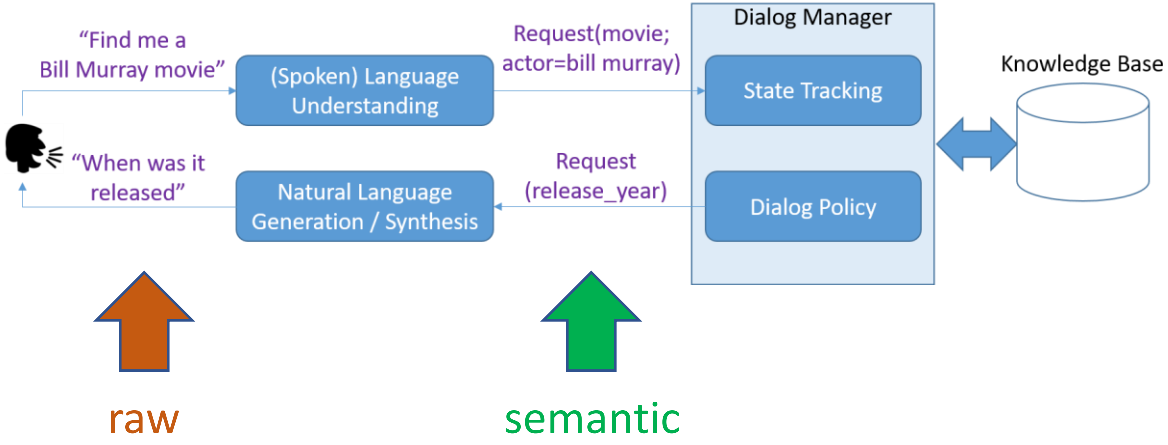


……

"Reinforcement Learning: An Introduction", 2nd ed., Sutton & Barto

Conversation as RL



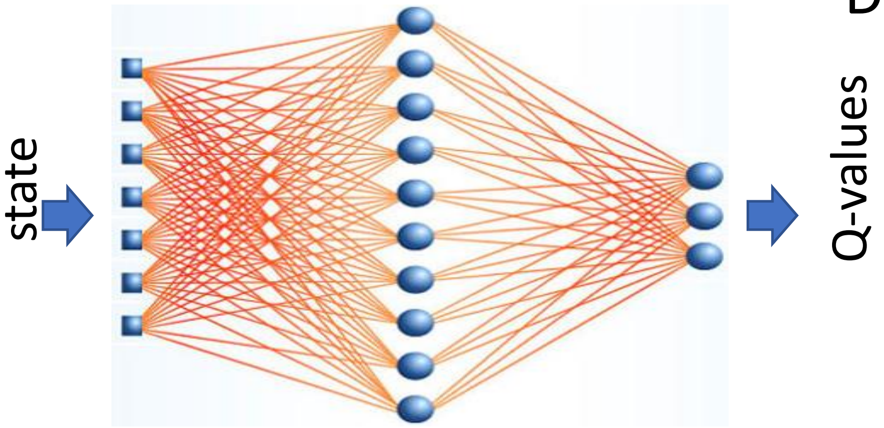


- State and action
  - Raw representation
    (utterances in natural language form)
  - Semantic representation
    (intent-slot-value form)
- Reward
  - +10 upon successful termination
  - -10 upon unsuccessful termination
  - -1 per turn
  - …

Pioneered by [Levin+ 00]

Other early examples: [Singh+ 02; Pietquin+ 04; Williams&Young 07; etc.]

Policy Optimization with DQN



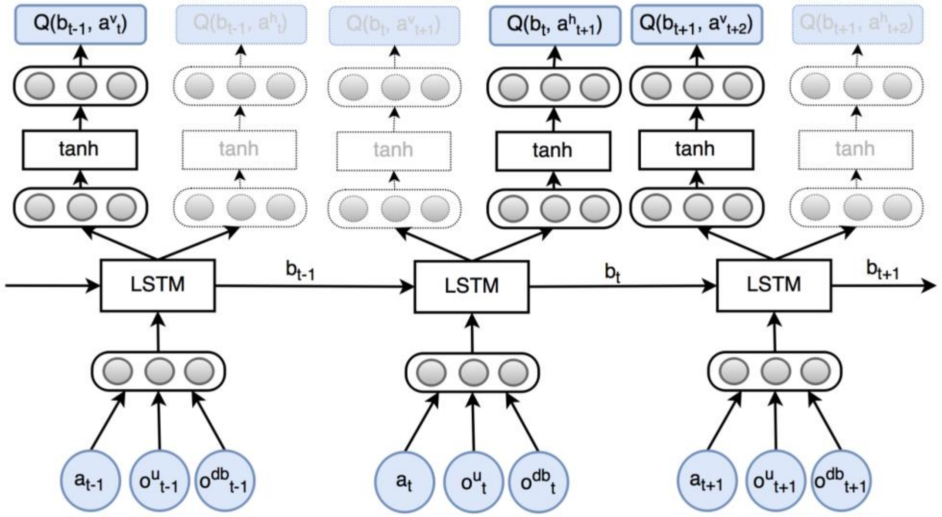
DQN-learning of network weights θ: apply SGD to solve

$$\hat{\theta} \leftarrow \arg\min_{\theta} \sum_{t} \left(r_{t+1} + \gamma \max_{a} Q_T(s_{t+1}, a) - Q_L(s_t, a_t) \right)^2$$

"Target network" to synthesize regression target

"Learning network" whose weights are to be updated

[Mnih+ 15]



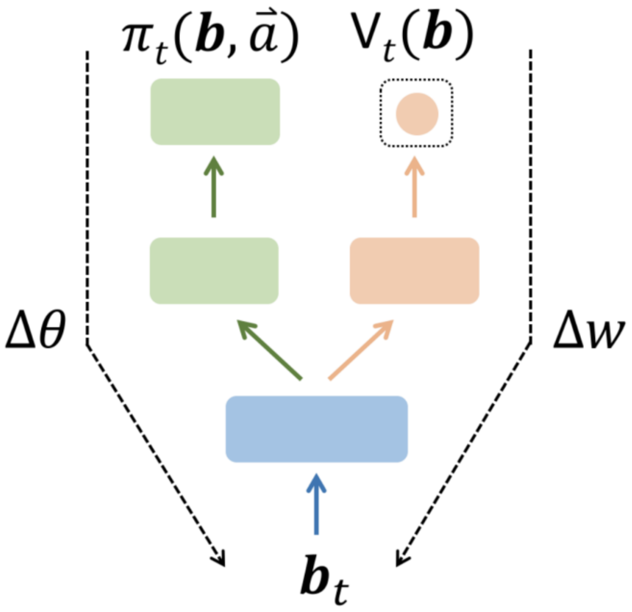
RNN/LSTM may be used to implicitly track states (without a separate dialogue state tracker)
[Zhao & Eskenazi 16]

Policy Optimization with Policy Gradient (PG)

- PG does gradient descent in policy parameter space to improve reward

$$\nabla_\theta J(\theta) = \mathbb{E}\left[\nabla_\theta \log \pi_\theta(a|\mathbf{b}) Q^{\pi_\theta}(\mathbf{b}, a)\right]$$

- REINFORCE [Williams 1992]: simplest PG algorithm
- Advantaged Actor-Critic (A2C) / TRACER
  - w: updated by least-squared regression
  - θ: updated as in PG



A2C/TRACER [Su+ 17]

Policy Gradient vs. Q-learning

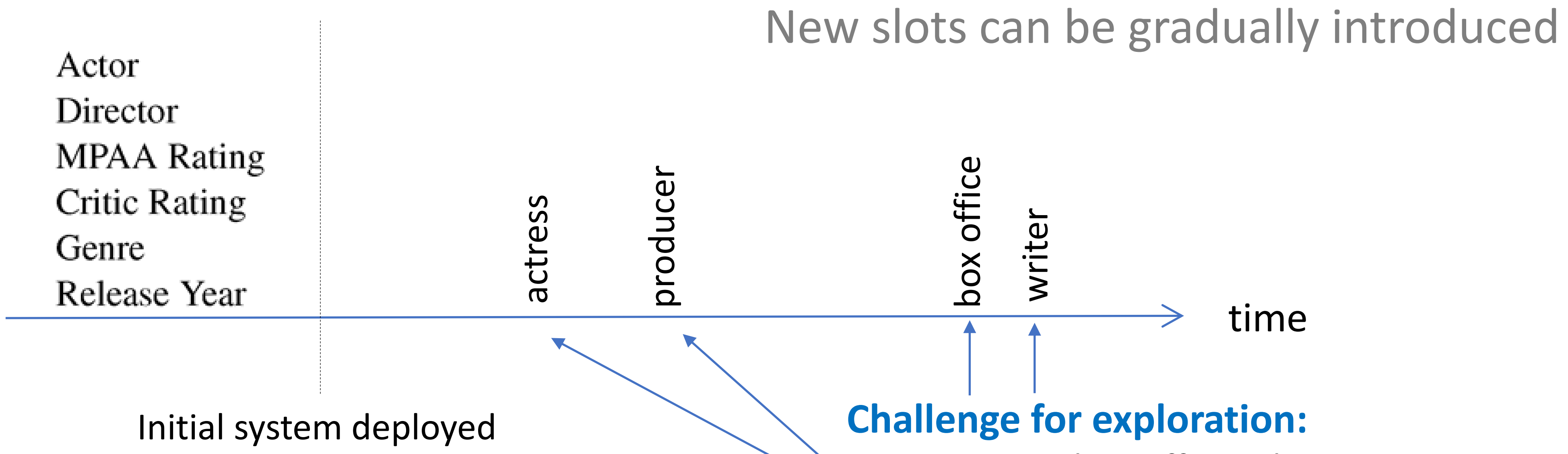
	Policy Gradient	**Q-learning**
Apply to complex actions	✓	
Stable convergence	✓	
Sample efficiency		✓
Relation to final policy quality	✓	
Flexibility in algorithmic design		✓

Three case studies

- How to efficiently explore the state-action space?
  - Modeling model uncertainty
- How to decompose complex state-action space?
  - Using hierarchical RL
- How to integrate planning into policy learning?
  - Balance the use of simulated and real experience – combining machine learning and machine teaching

Domain Extension and Exploration

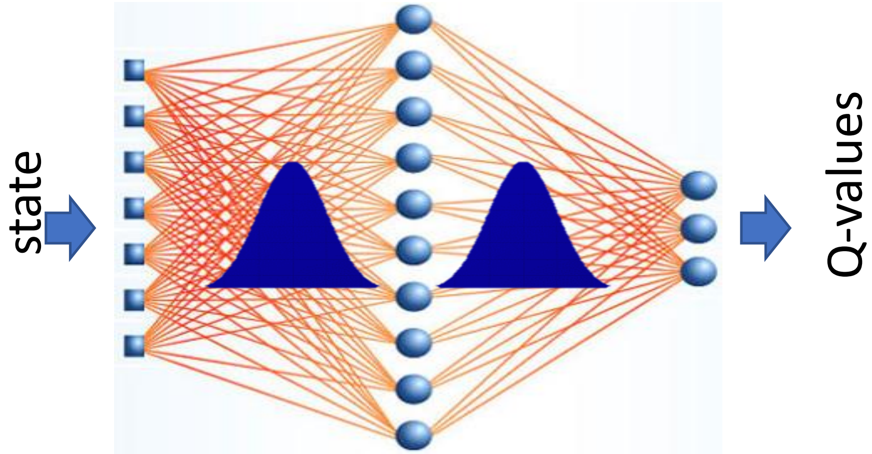
- Most goal-oriented dialogs require a closed and well-defined domain
- Hard to include all domain-specific information up-front

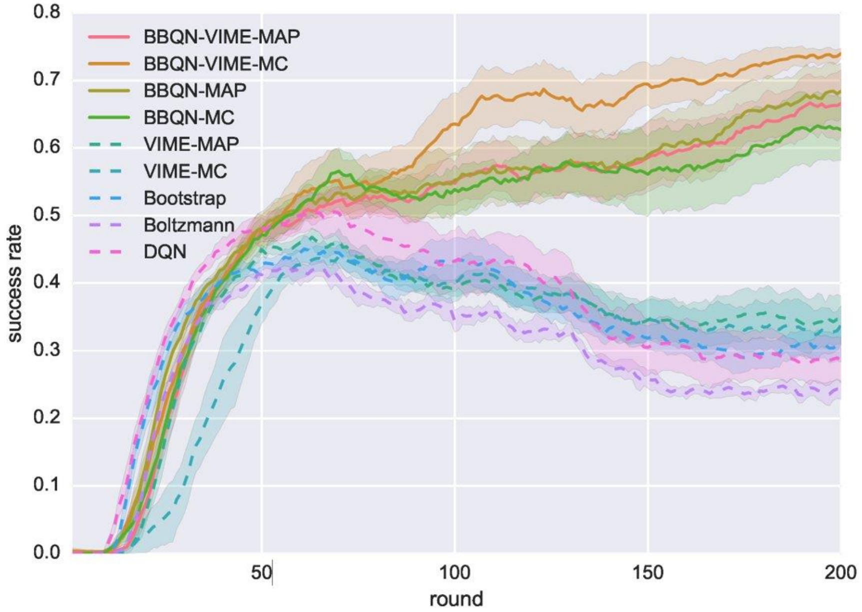


Challenge for exploration:

- How to explore efficiently
- to collect data for new slots
- When deep models are used

Bayes-by-Backprop Q (BBQ) network





BBQ-learning of network params $\theta = (\mu, \sigma^2)$:

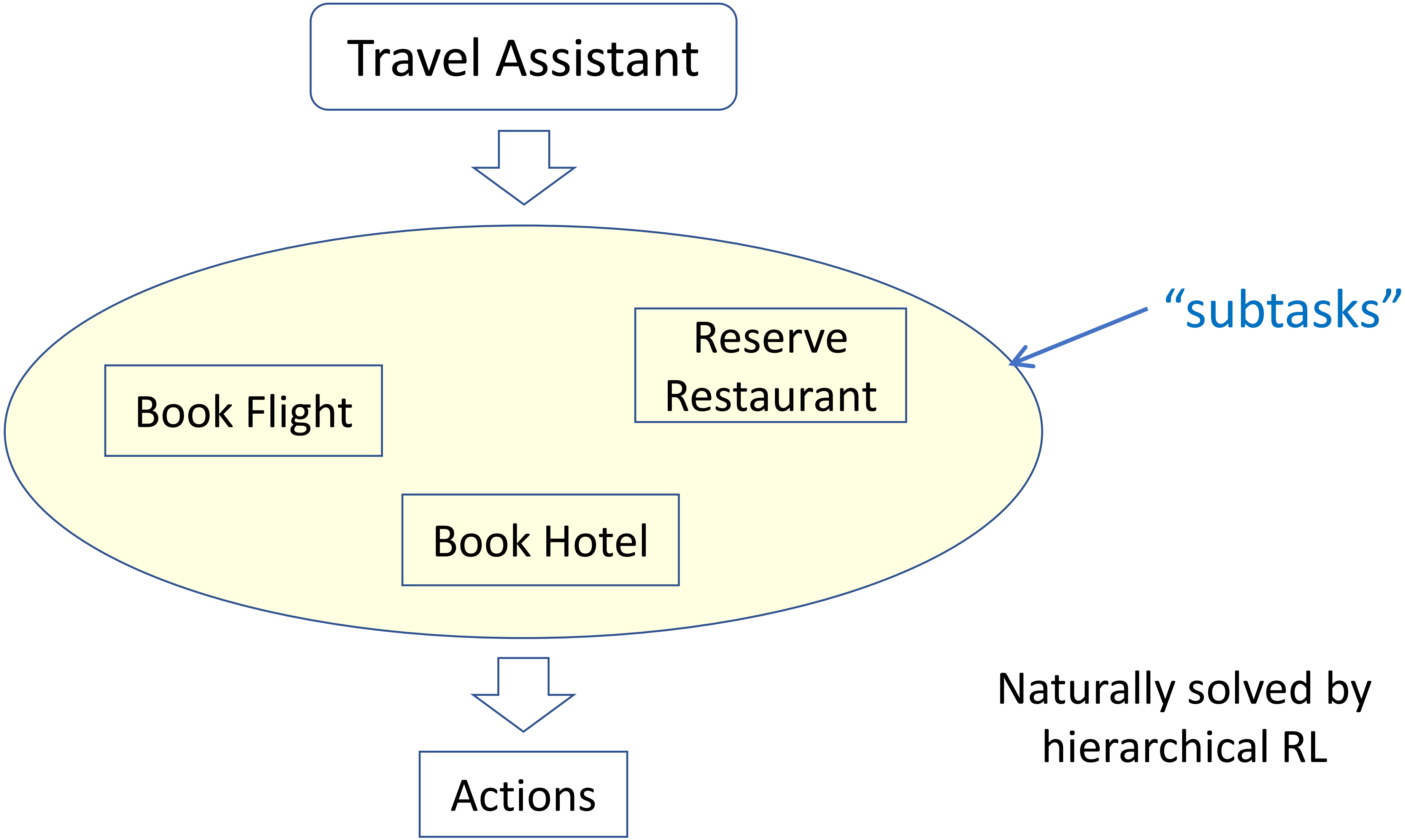
$$\hat{\theta} = \arg\min_{\theta_L} \mathrm{KL}\big(q(\mathbf{w}|\theta_L) \;||\; p(\mathbf{w}|Data)\big)$$

Still use "target network" θ_T
to synthesize regression target

- Parameter learning: solve for $\hat{\theta}$ with Bayes-by-backprop [Blundell et al. 2015]
- Params θ quantifies uncertainty in Q-values
- Action selection: use Thompson sampling for exploration

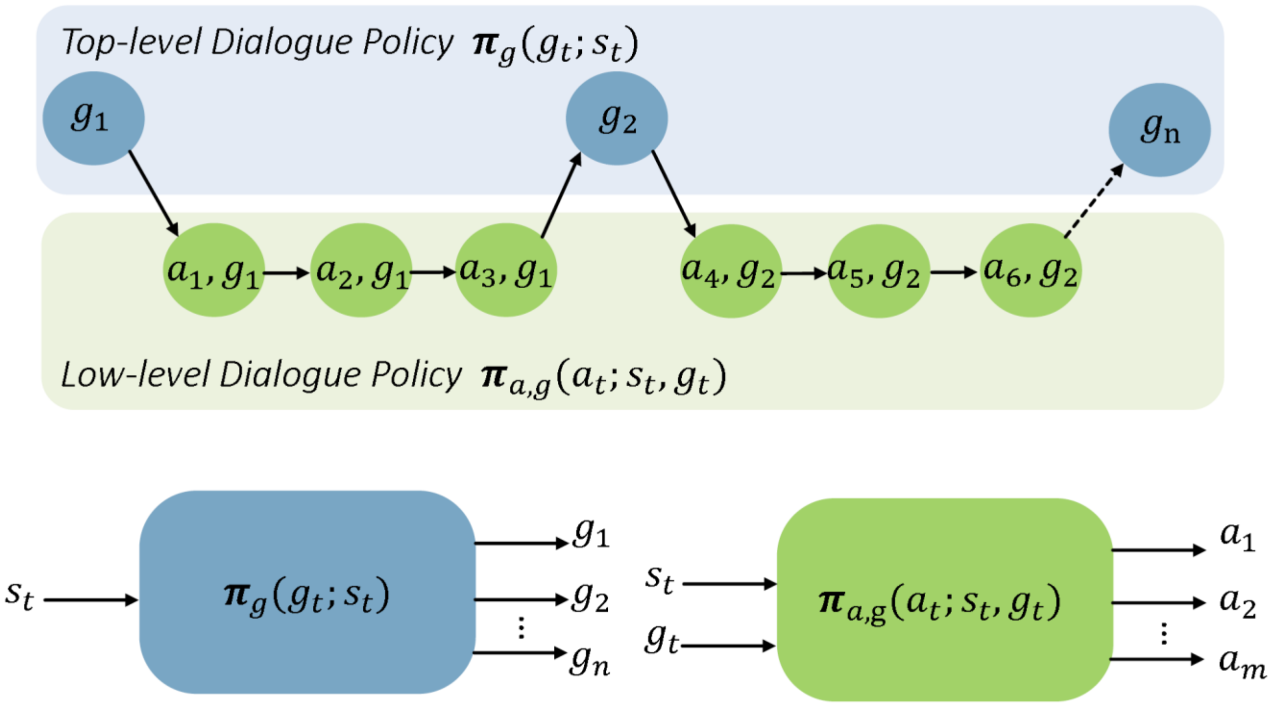
[Lipton+ 18]

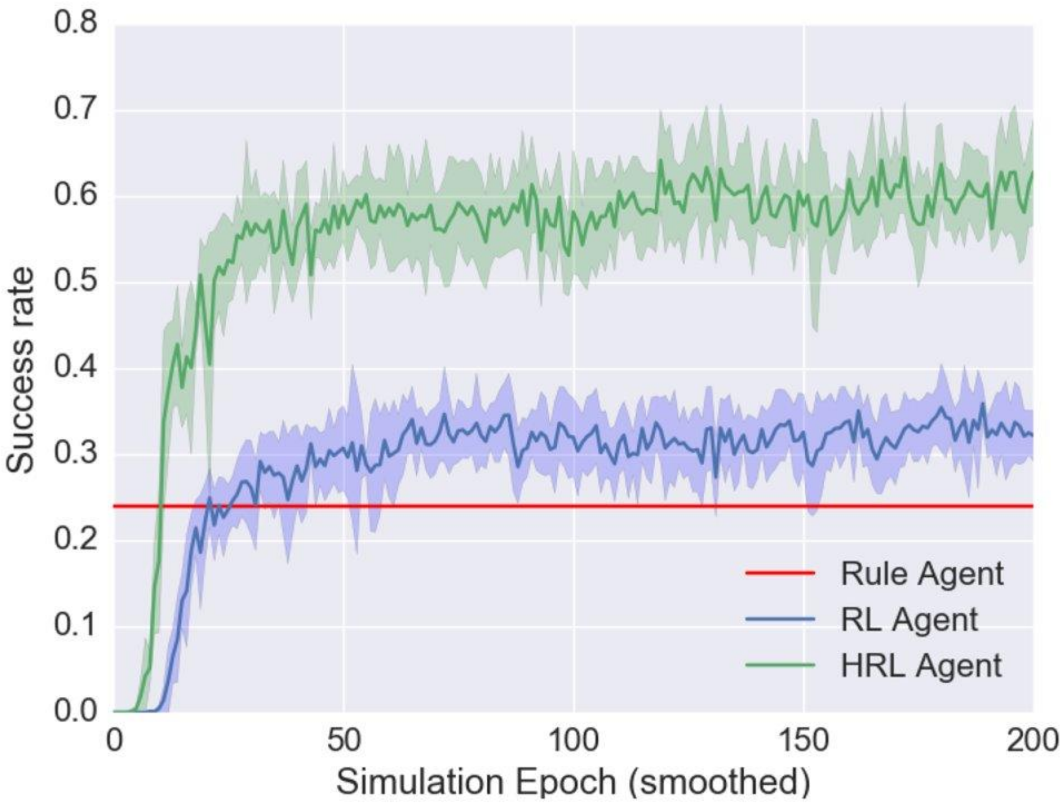
Composite-task Dialogues



Naturally solved by hierarchical RL

A Hierarchical Policy Learner



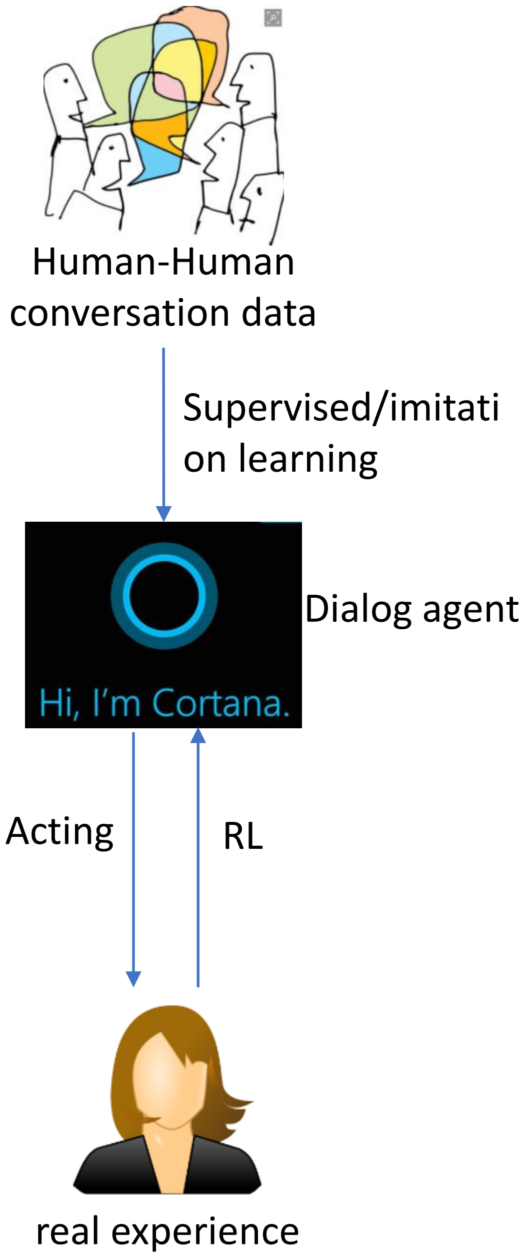


Similar to Hierarchical Abstract Machine (HAM) [Parr'98]

Superior results in both simulated and real users [Peng+ 17]

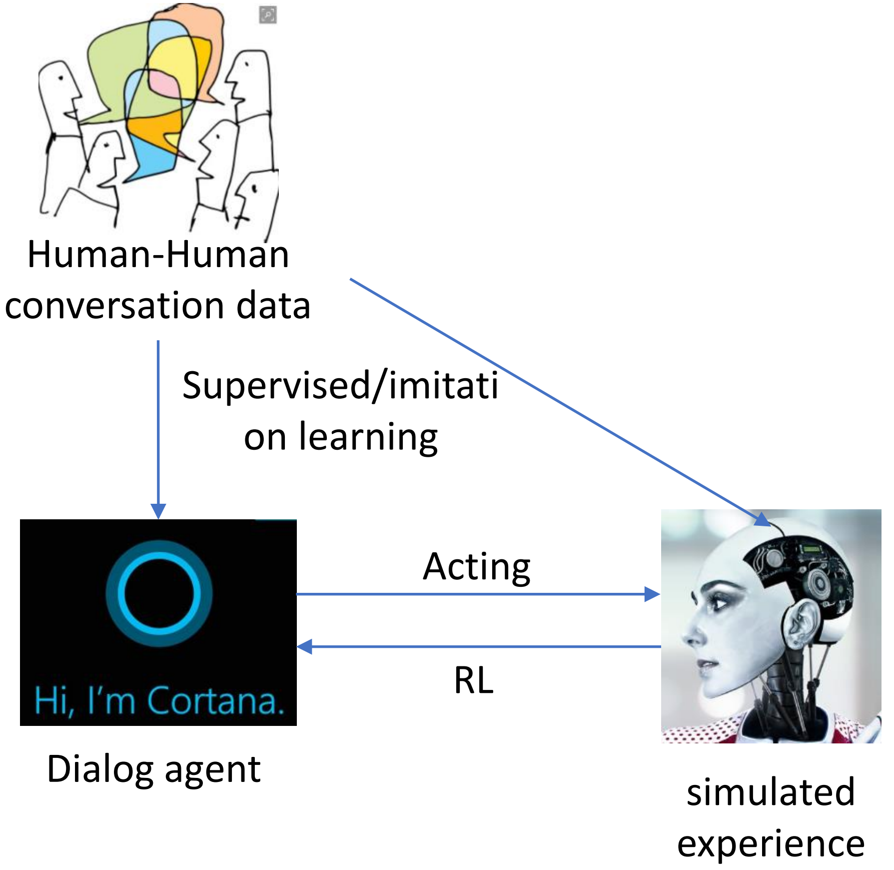
Integrating Planning for Dialogue Policy Learning [Peng+ 18]

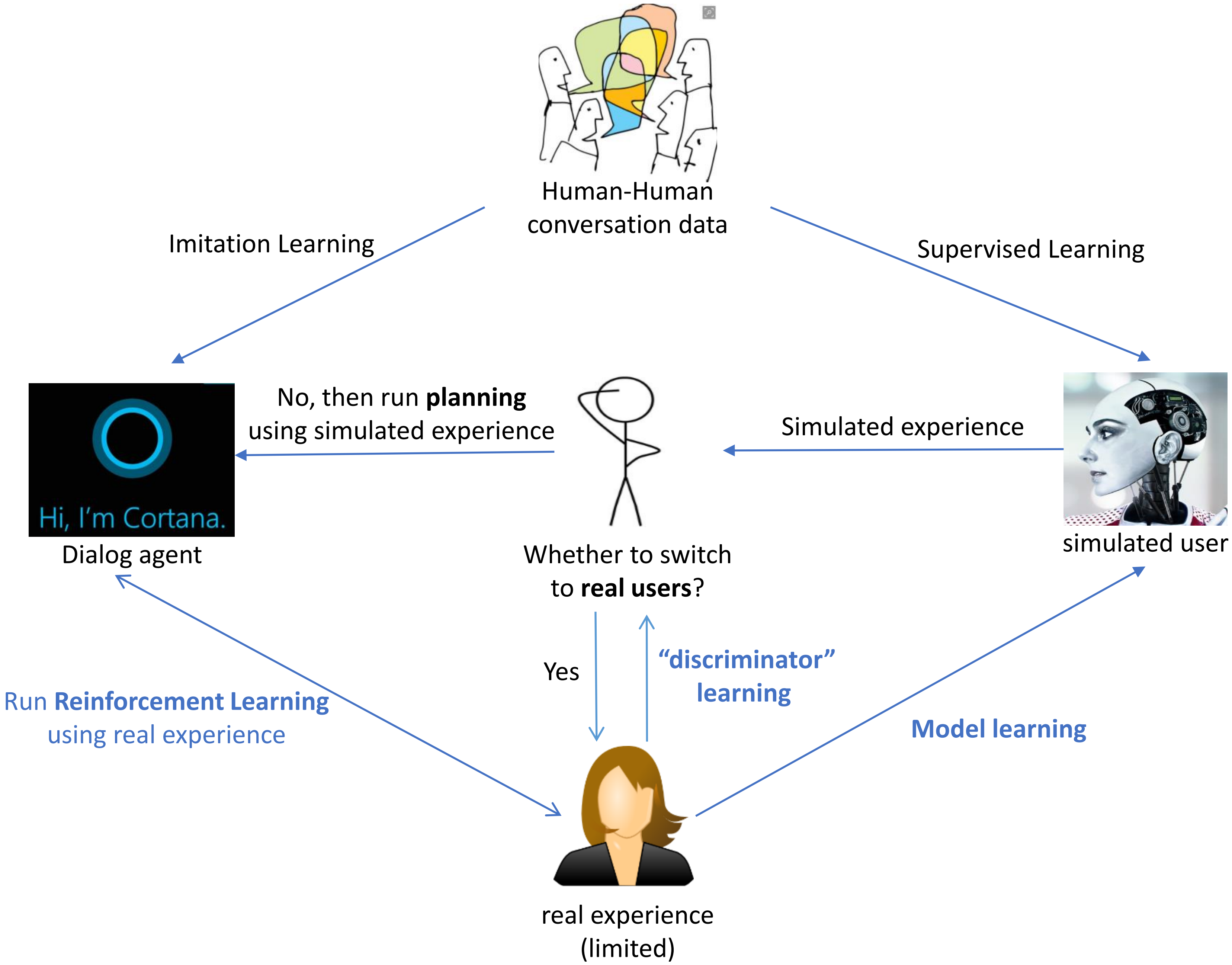
- Expensive: need large amounts of real experience except for very simple tasks
- Risky: bad experiences (during exploration) drive users away



Integrating Planning for Dialogue Policy Learning [Peng+ 18]

- Inexpensive: generate large amounts of simulated experience for free
- Overfitting: discrepancy btw real users and simulators





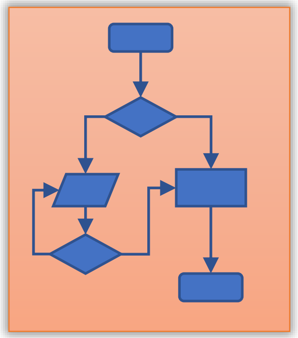
[Peng+ 18, Su +18, Wu + 19, Zhang+ 19,]

Programmatic	Declarative	Machine Learning
▼ Accessible to non-experts	▲ Accessible to non-experts	▲ Accessible to non-experts
▼ Easy to debug	▼ Easy to debug	▲ Easy to debug
▲ Explicit Control	▲ Explicit Control	▼ Explicit Control
▲ Support for complex scenarios	▼ Support for complex scenarios	▲ Support for complex scenarios
▼ Ease of Modification	▼ Ease of Modification	▲ Ease of Modification
▼ Handle Unexpected Input	▼ Handle Unexpected Input	▼ Handle Unexpected Input
▼ Improve / Learn from conversations	▼ Improve / Learn from conversations	▲ Improve / Learn from conversations
▲ No Dialog Data Required	▲ No Dialog Data Required	▼ Requires Sample Dialog Data

Programmatic

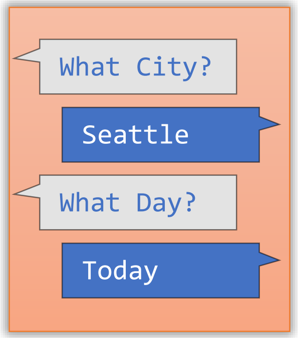
```
this.dialogs.add(
 new WaterfallDialog(GET_FORM_DATA,
  [
   this.askForCity.bind(this),
   this.collectAndDisplayName.bind(this)
  ]
));
async collectAndDisplayName(step) {…
```

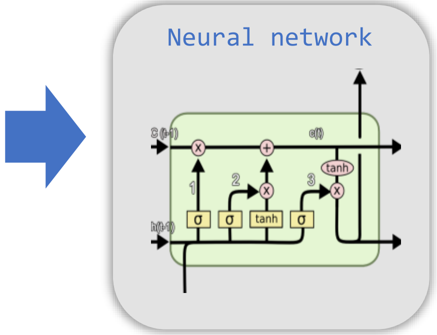
Declarative



```
<rule>
  <if>
    city == null
  </if>
  <then>
    Which city?
  </then>
   …
```

Machine Learning

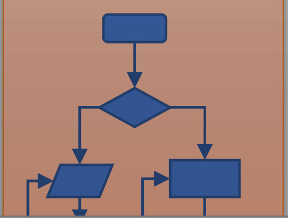




Microsoft

Programmatic	Declarative	Machine Learning

```
this.dialogs.add(
 new WaterfallDialog(GET_FORM_DATA,
  [
   this.askForCity.bind(this),
   this.collectAndDisplayName.bind(this)
  ]
));
async collect
```

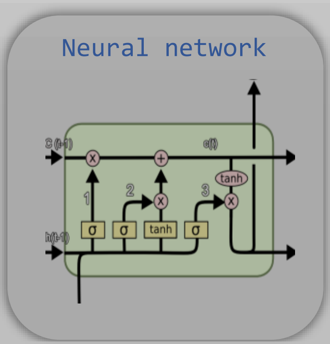


```
<rule>
  <if>
    city == null
  </if>
  <then>
    Which city?
```

What City?

Seattle

What Day?



One Solution Does Not Fit All

- Accessible
- Easy to de
- Explicit Co
- Support fo
- Ease of M
- Handle Unexpected Input
- Improve / Learn from conversations
- No Dialog Data Required

- Handle Unexpected Input
- Improve / Learn from conversations
- No Dialog Data Required

- on-experts
- mplex scenarios
- cation
- Handle Unexpected Input
- Improve / Learn from conversations
- Requires Sample Dialog Data

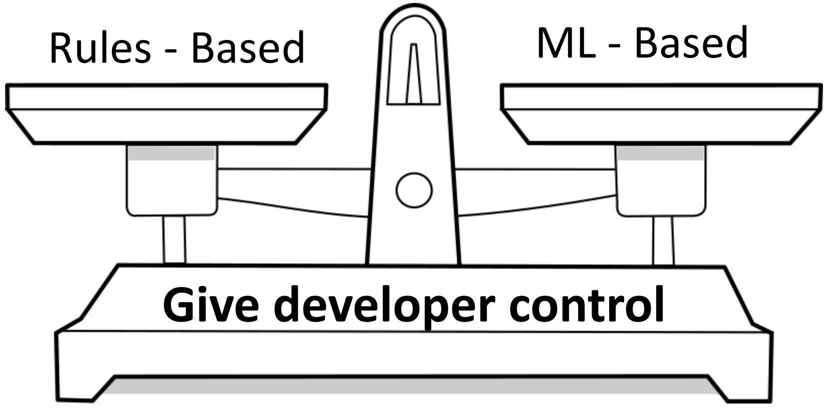
Goal: Best of both worlds

Good for garden path

Not data intensive

Explicit Control

Easily interpretable



Handle unexpected input

Learn from usage data

Often viewed as black box

Start with rules-based policy => Grow with Machine Learning

Make ML more controllable by visualization

Not unidirectional : Rules-based policy can evolve side-by-side with ML Model

Conversation Learner – building a bot interactively

What is it: A system built on the principles of Machine Teaching, that enables individuals with no AI experience (designers, business owners) to build task-oriented conversational bots

Goal: Push the forefront of research on conversational systems using input from enterprise customers and product teams to provide grounded direction for research

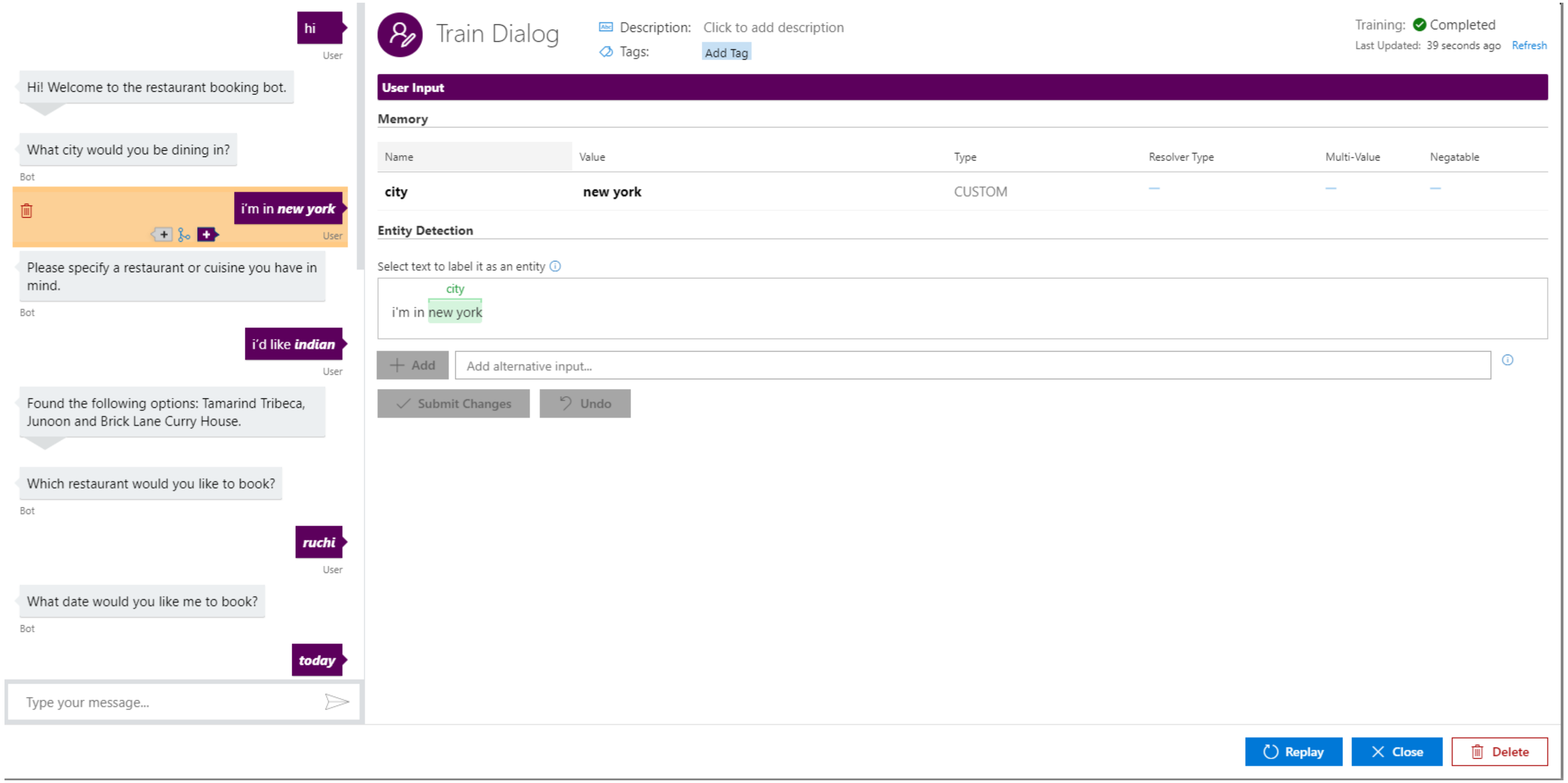
Status: In private preview with ~50 customers to various levels of prototyping

Hello World Tutorial

Primary repository with samples:
https://github.com/Microsoft/ConversationLearner-samples

Conversation Learner – building a bot interactively

- Rich machine teaching and dialog management interface accessible to non-experts
- Free-form tagging, editing and working directly with conversations
- Incorporating rules makes the teaching go faster
- Independent authoring of examples allows dialog authors to collaborate on one/multiple intents



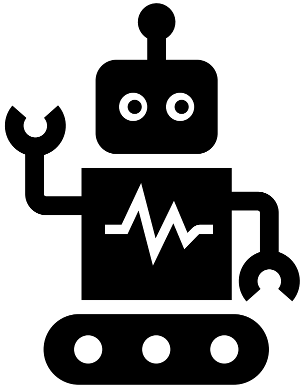
ConvLab

[Published @ https://arxiv.org/abs/1904.08637](https://arxiv.org/abs/1904.08637)

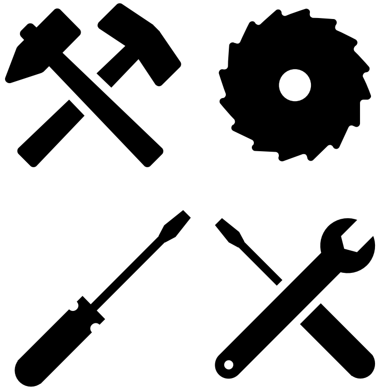
Fully annotate data
for training individual components or end-to-end models with supervision

Speaker	Utterance	Annotation
User	am looking for a place to to stay that has cheap price range it should be in a type of hotel	**Dialog acts** {"Hotel-Inform": [["Pricerange", "cheap"]]}
System	Okay, do you have a specific area you want to stay in?	**State** {"hotel": { "name": "not mentioned", "area": "not mentioned", "parking": "not mentioned", "pricerange": "cheap", "stars": "not mentioned", "internet": "not mentioned", "type": "hotel"}}} ... **Dialog acts** "Hotel-Request": [["Area", "?"]]
User	no, i just need to make sure it's cheap. oh, and i need parking	**Dialog acts** {"negate", "Hotel-Inform": [["Pricerange", "cheap"], ["Parking", "yes"]]}
System	I found 1 cheap hotel for you that includes parking. Do you like me to book it?	**State** {"hotel": { "name": "not mentioned", "area": "not mentioned", "parking": "yes", "pricerange": "cheap", "stars": "not mentioned", "internet": "not mentioned", "type": "hotel"}}} ... **Dialog acts** {"Hotel-Inform": [["Price", "cheap"], ["Choice", "1"], ["Parking", "none"]]}

User Simulators
for reinforcement learning
1 rule-based simulator
2 data-driven simulators



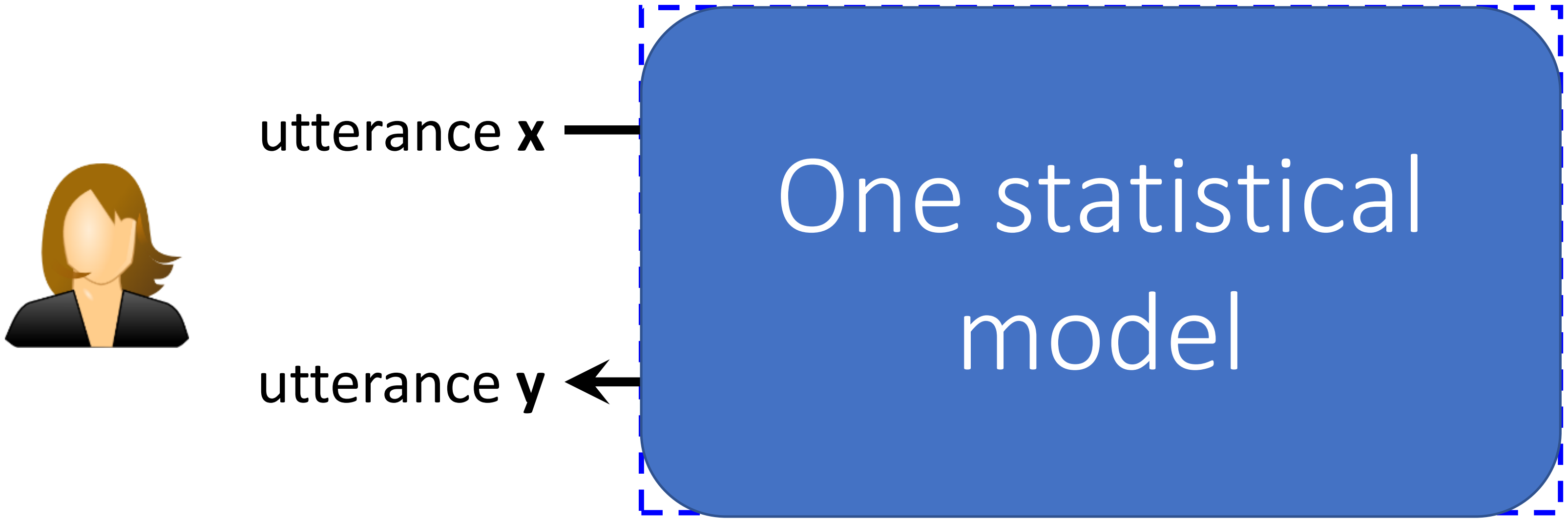
SOTA Baselines
Multiple models for each component
Multiple end-to-end system recipes



Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogue
- **Part 4: Fully data-driven conversation models and chatbots**
  - **E2E neural conversation models**
  - **Challenges and remedies**
  - **Grounded conversation models**
  - **Beyond supervised learning**
  - **Data and evaluation**
  - **Chatbots in public**
  - **Future work**

Motivation



Move towards **fully data-driven**, **end-to-end** dialogue systems.

Social Bots

- Fully end-to-end systems so far most successfully applied to **social bots** or **chatbots**:
  - Commercial systems: Amazon Alexa, XiaoIce, etc.

- Why social bots?
  - Maximize **user engagement** by generating **enjoyable** and **more human-like** conversations
  - Help **reduce user frustration**
  - **Influence dialogue research** in general (social bot papers often cited in task-completion dialogue papers)

Historical overview

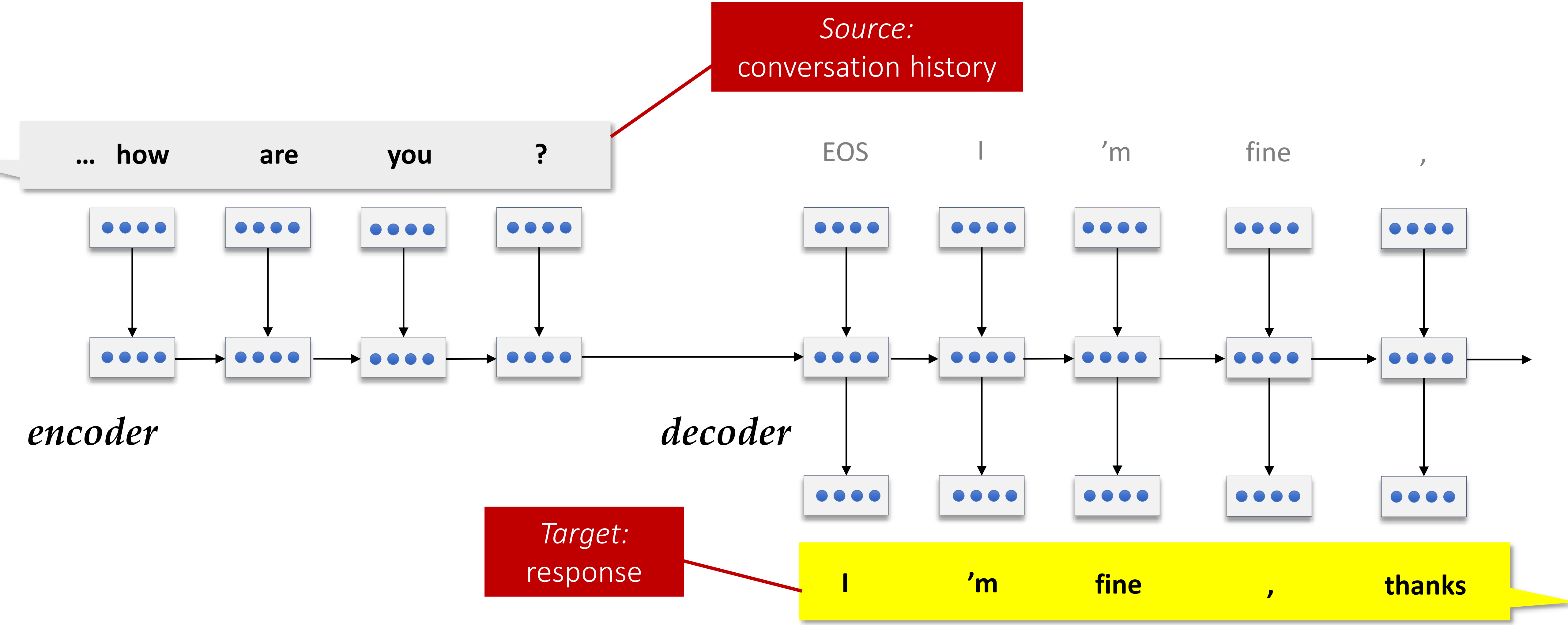
Earlier work in fully data-driven response generation:

- **2010:** Response retrieval system (IR) [Jafarpour+ 10]
- **2011**: Response generation using Statistical Machine Translation (phrase-based MT) [Ritter+ 11]
- **2015**: First neural response generation systems (RNN, seq2seq)

  [Sordoni+ 15; Vinyals & Le 15; Shang+ 15]

Neural Models for Response Generation

[Sordoni+ 15; Vinyals & Le 15; Shang+ 15]



Similar to sequence models in Neural Machine Translation (NMT), summarization, etc.
Uses either RNN, LSTM, GRU, Pointer-Generator Networks, Transformer, etc.

Neural Response Generation:
Difference with other tasks (e.g., machine translation)

- **Data:** some training sets (social media) are **HUGE**

  For example, Twitter (as of 2016):

  304M monthly active users

  500M tweets per day (6M conversations per day)

→ *often train on subset of the data; leaner and "faster" models preferred*

- **System input:** LONG conversation history

$T_1\ T_2\ T_3\ T_4 \ldots T_N$

Conversation (N turns)

$T_1 \longrightarrow T_2$

$T_1\ T_2 \longrightarrow T_3$

…

$\ldots\ T_{N-2}\ T_{N-1} \longrightarrow T_N$

N-1 training instances

Either limit context size [Sordoni+ 05] or truncate backprop [Vinyals+ 05, pc]

[Sordoni+ 15; Li+ 16a]

Neural conversation engine

Neural model learns the general shape of conversation, and the system output is situationally appropriate and coherent…

H:	Hi
M:	**How are you?**
H:	Not bad
M:	**What's wrong?**
H:	Nothing really
M:	**Wanna talk about it? I'm here if you wanna talk x**
H:	Well, I'm just tired
M:	**Me too, I'm here if you wanna talk**
H:	Thanks, what are you doing?
M:	**You're welcome and I'm watching TV and you?**
H:	On Netflix?
M:	**Yeah, I'm watching it on Netflix.**
H:	What are you watching?
M:	**A TV show.**
H:	Nice, did you have dinner already?
M:	**No, not yet. what about you?**

A more goal-oriented example. No grounding into a real calendar, but the “shape” of the conversation is fluent and plausible…

H:	Would Thursday afternoon work sometime?
M:	**Yeah, sure. just let me know when you‘re free.**
H:	After lunch is probably the best time
M:	**Okay, sounds good. Just let me know when you’re free.**
H:	Would 2pm work for you?
M:	**Works for me.**
H:	Well let’s say 2pm then I’ll see you there
M:	**Sounds good.**

Neural E2E system trained on 35M Twitter conversations.

Fully Data-driven Response Generation:

Challenges and remedies

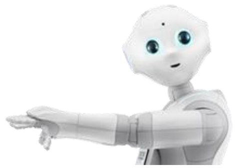
Challenge: The blandness problem

How was your weekend?

I don't know.

What did you do?

I don't understand what you are talking about.



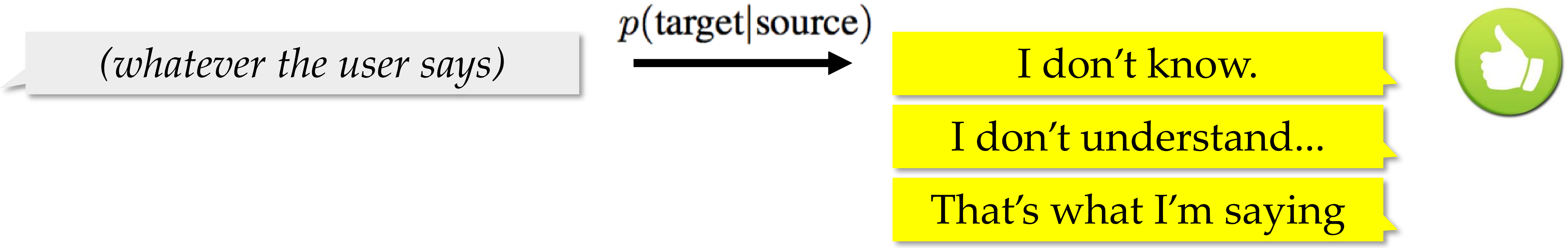
This is getting boring…

Yes that's what I'm saying.

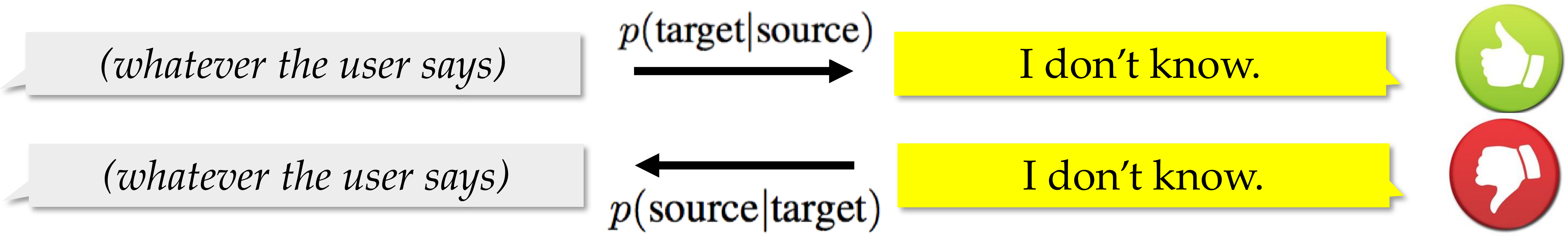
[Li+ 16a]

Blandness problem: cause and remedies

Common MLE objective (maximum likelihood)



Mutual information objective:



[Li+ 16a]

Mutual Information for Neural Network Generation

Mutual information objective:

$$\hat{T} = \arg\max_T \left\{ \log \frac{p(S,T)}{p(S)p(T)} \right\}$$

$$\hat{T} = \arg\max_T \left\{ \underbrace{\boxed{\log p(T|S)}}_{\text{standard likelihood}} - \underbrace{\boxed{\lambda \log p(T)}}_{\text{anti-LM}} \right\}$$

$$p(\text{target}|\text{source}) \longrightarrow \qquad \longleftarrow p(\text{source}|\text{target})$$

Bayes' rule

Bayes' theorem

$$\hat{T} = \arg\max_T \left\{ (1-\lambda) \log p(T|S) + \lambda \log p(S|T) \right\}$$

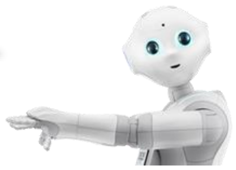
Sample outputs (MMI)

Wow sour starbursts really do make your mouth water... mm drool. Can I have one?

Of course you can! They're delicious!

Milan apparently selling Zlatan to balance the books... Where next, Madrid?

I think he'd be a good signing.



'tis a fine brew on a day like this! Strong though, how many is sensible?

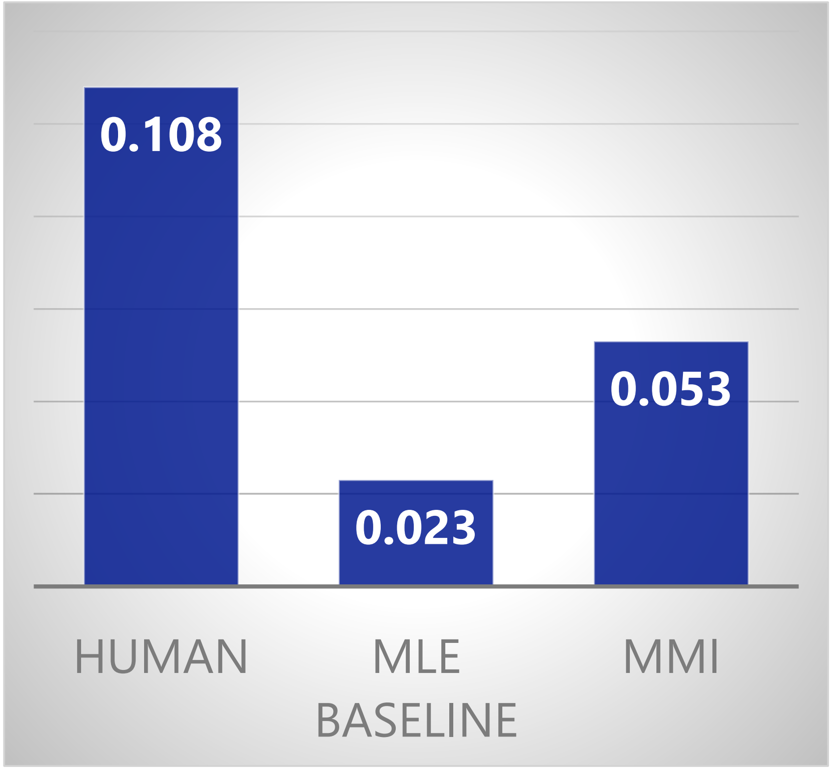
Depends on how much you drink!

Well he was on in Bromley a while ago... still touring.

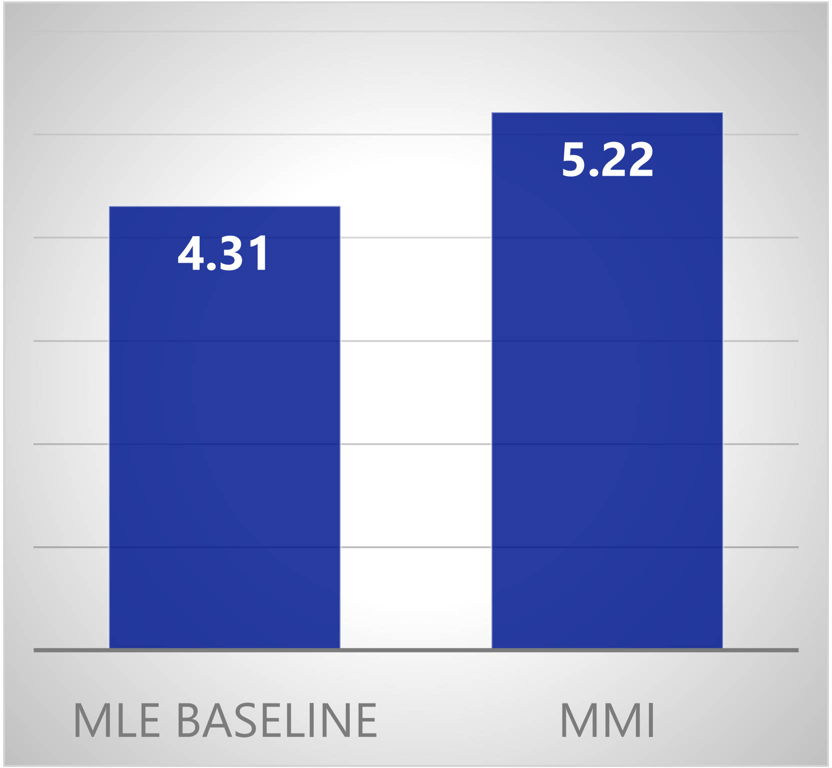
I've never seen him live.

[Li+ 16a]

MLE vs MMI: results



Lexical diversity
(# of distinct tokens / # of words)



BLEU

MMI: best system in Dialogue Systems Technology Challenge 2017 (**DSTC**, E2E track)

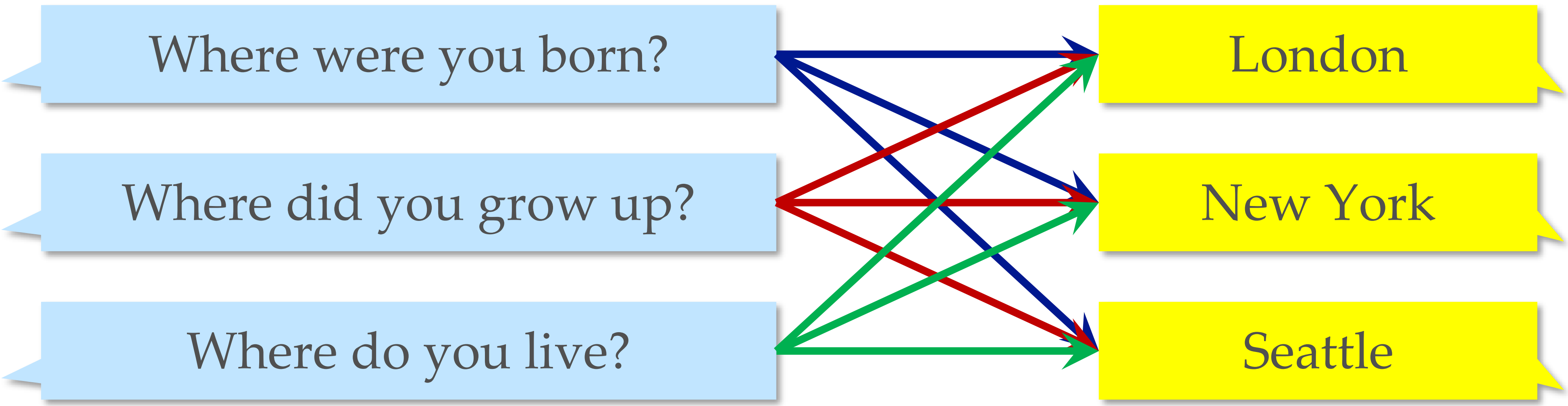
Challenge: The consistency problem

- E2E systems often exhibit **poor response consistency**:

message	Where is your hometown?
response	I live in Indonesia.
message	Where are you from?
response	England, you?
message	In which city do you live now?
response	I live in Jakarta.
message	In which country do you live now?
response	I live in Indonesia.
message	What is your major?
response	Business, you?
message	What did you study in college?
response	Psychology, you?
message	How old are you?
response	I'm 18.
message	What is your age?
response	I'm 16.

The consistency problem: why?

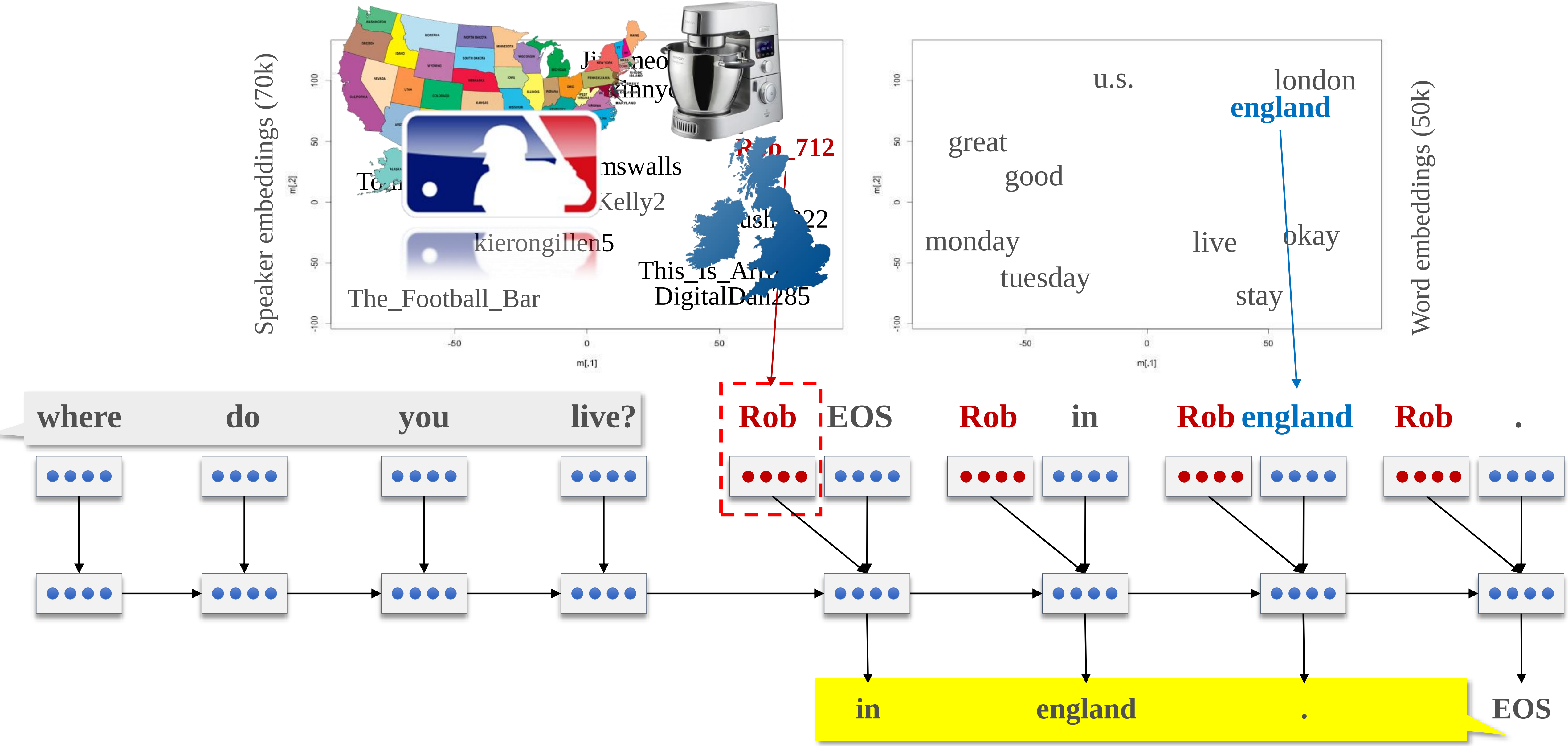
Conversational data:



NOT 1-to-1

P(response | query, **SPEAKER_ID**)

Personalized Response Generation [Li+ 2016b]



Persona model results

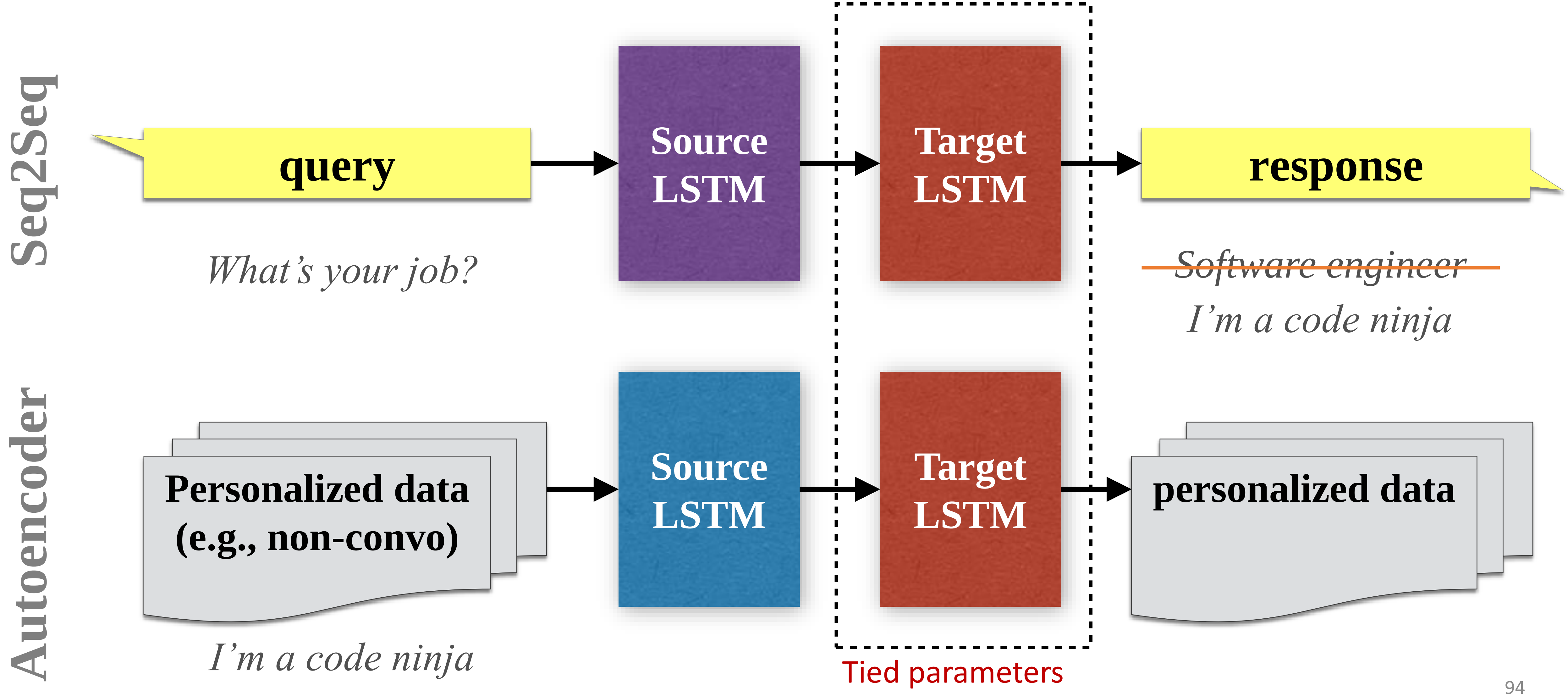
Baseline model:

message	Where is your hometown?
response	I live in Indonesia.
message	Where are you from?
response	England, you?
message	In which city do you live now?
response	I live in Jakarta.
message	In which country do you live now?
response	I live in Indonesia.
message	What is your major?
response	Business, you?
message	What did you study in college?
response	Psychology, you?
message	How old are you?
response	I'm 18.
message	What is your age?
response	I'm 16.

Persona model using speaker embedding: [Li+ 16b]

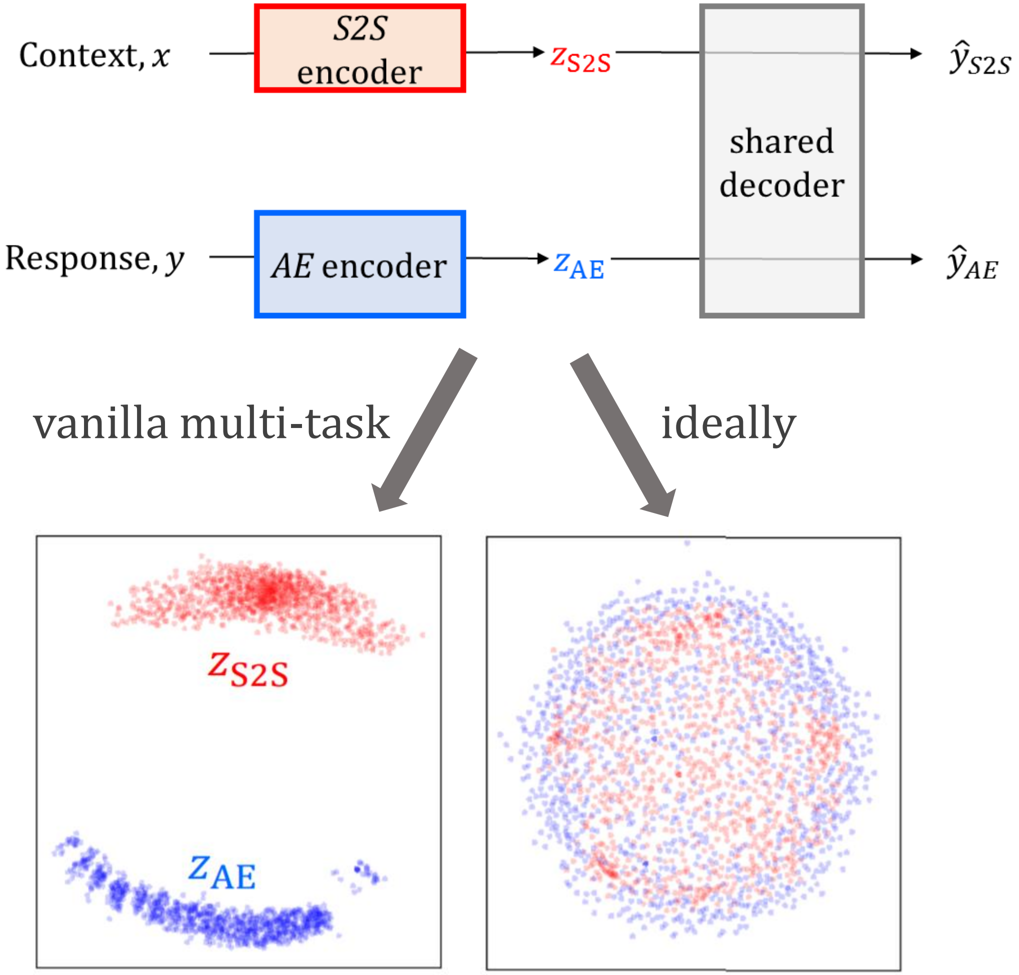
message	Where is your hometown?
response	I'm from England.
message	Where are you from?
response	I'm from England.
message	In which city do you live now?
response	I live in London.
message	In which country do you live now?
response	I live in England.
message	What is your major?
response	Business. You?
message	What did you study in college?
response	I did business studies.
message	How old are you?
response	I'm 18.
message	What is your age?
response	I'm 18.

Personal modeling as multi-task learning [Luan+ 17]



Challenges with multi-task learning

[Gao+ 19]



So we add regularization:

$$\mathcal{L} = \boxed{-\frac{1}{|y|}\log p(y|z_{\mathrm{S2S}}) - \frac{1}{|y|}\log p(y|z_{\mathrm{AE}})} + \alpha\mathcal{L}_{\mathrm{fuse}} + \ldots$$

Vanilla S2S + Mtask objective

where:

$$\mathcal{L}_{\mathrm{fuse}} = \boxed{\sum_{i\in\mathrm{batch}} \frac{d(z_{\mathrm{S2S}}(x_i), z_{\mathrm{AE}}(y_i))}{n}}$$

cross-space distance

$$- \boxed{\sum_{i,j\in\mathrm{batch}, i\neq j} \frac{d(z_{\mathrm{S2S}}(x_i), z_{\mathrm{S2S}}(x_j))}{n^2-n}} - \boxed{\sum_{i,j\in\mathrm{batch}, i\neq j} \frac{d(z_{\mathrm{AE}}(y_i), z_{\mathrm{AE}}(y_j))}{n^2-n}}$$

same-space distance

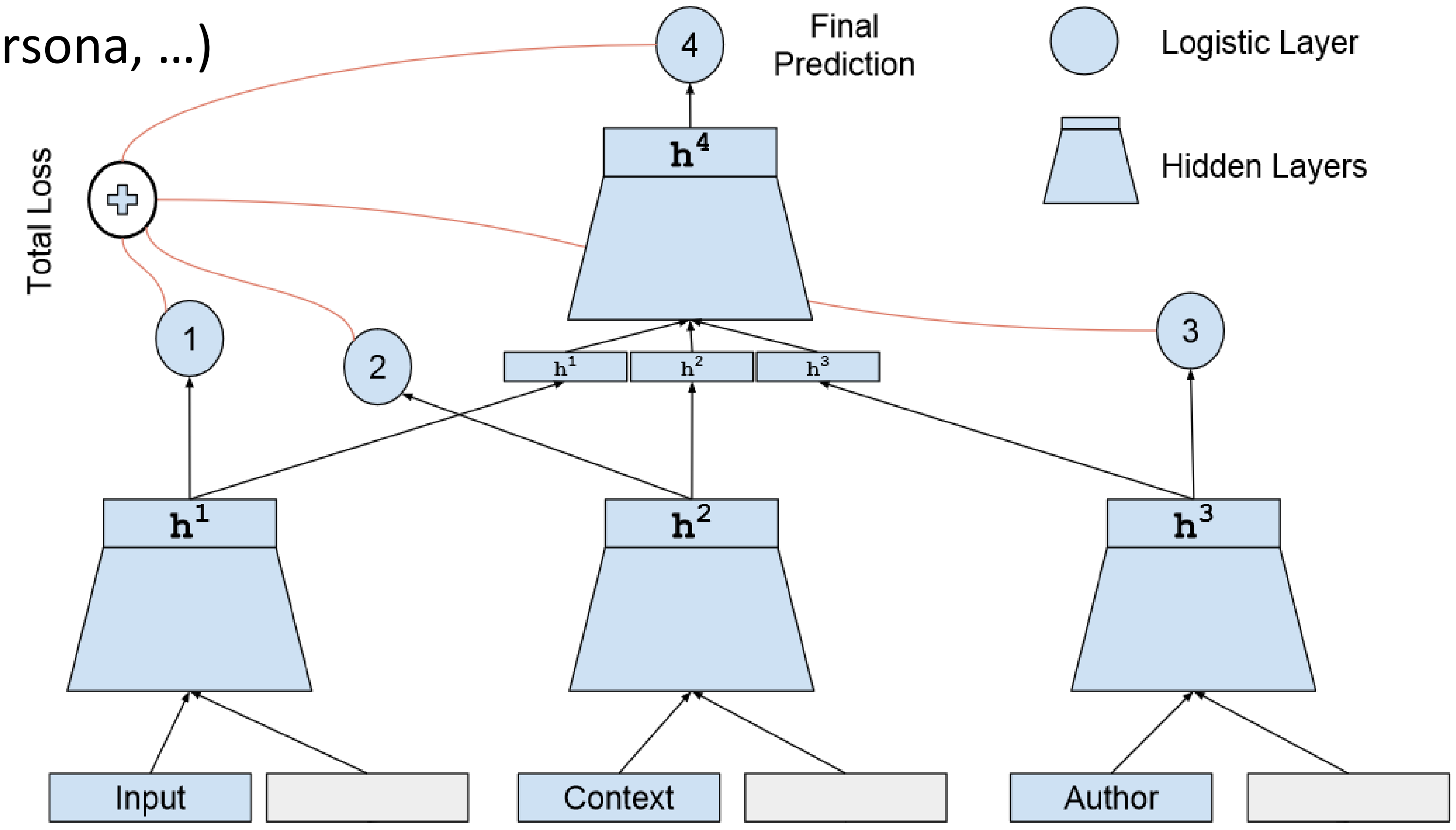
Improving personalization with multiple losses [Al-Rfou+ 16]

- **Single-loss:**
  P(response | context, query, persona, …)

  Problem with single-loss: context or query often **“explain away”** persona

- **Multiple loss adds:**
  P(response | persona)
  P(response | query)
  etc.

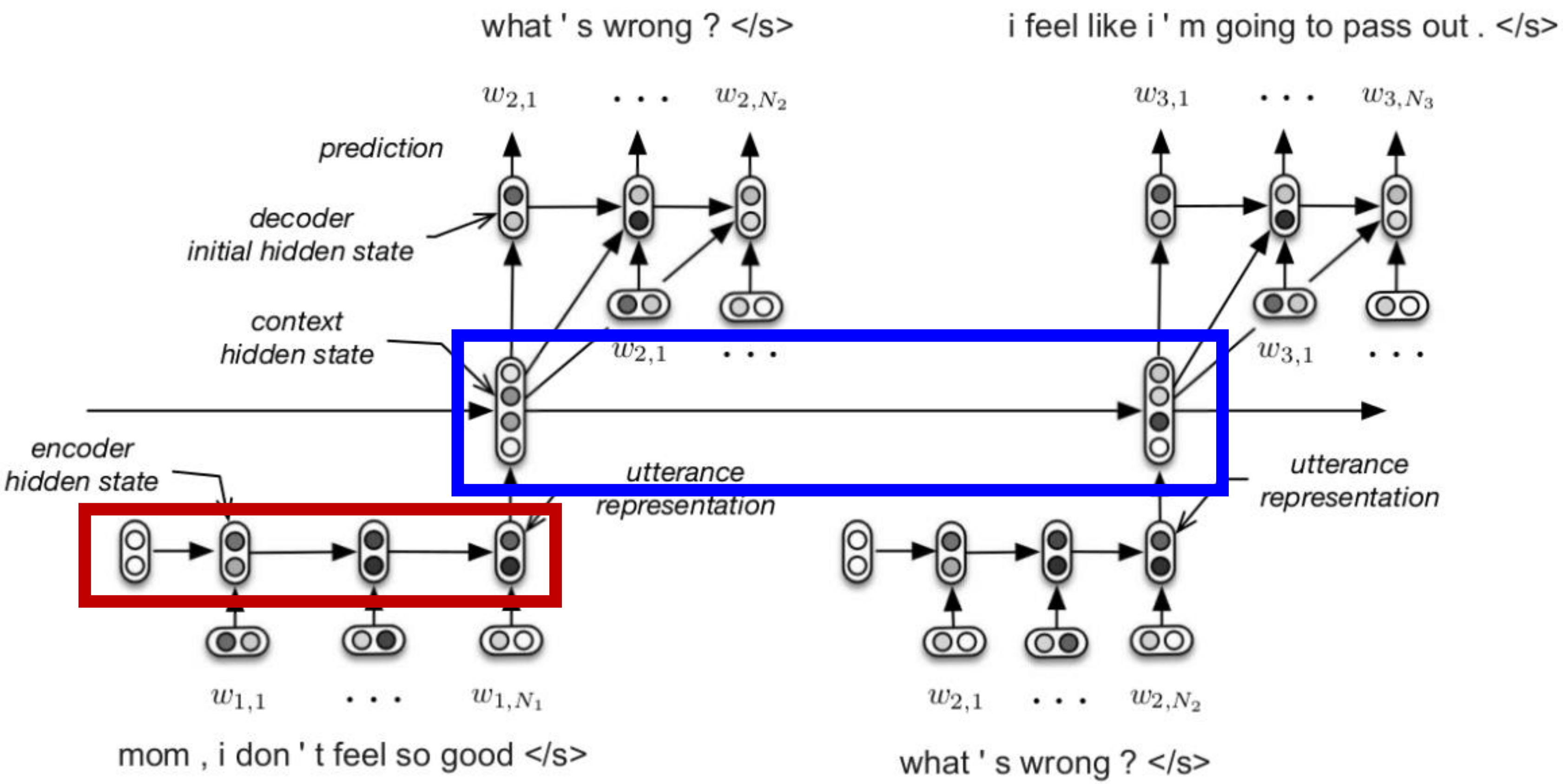
  Optimized so that persona can “predict” response all by itself
  → more robust speaker embeddings



Challenge: Long conversational context

It can be challenging for LSTM/GRU to encode very long context (i.e. more than 200 words: [Khandelwal+ 18])

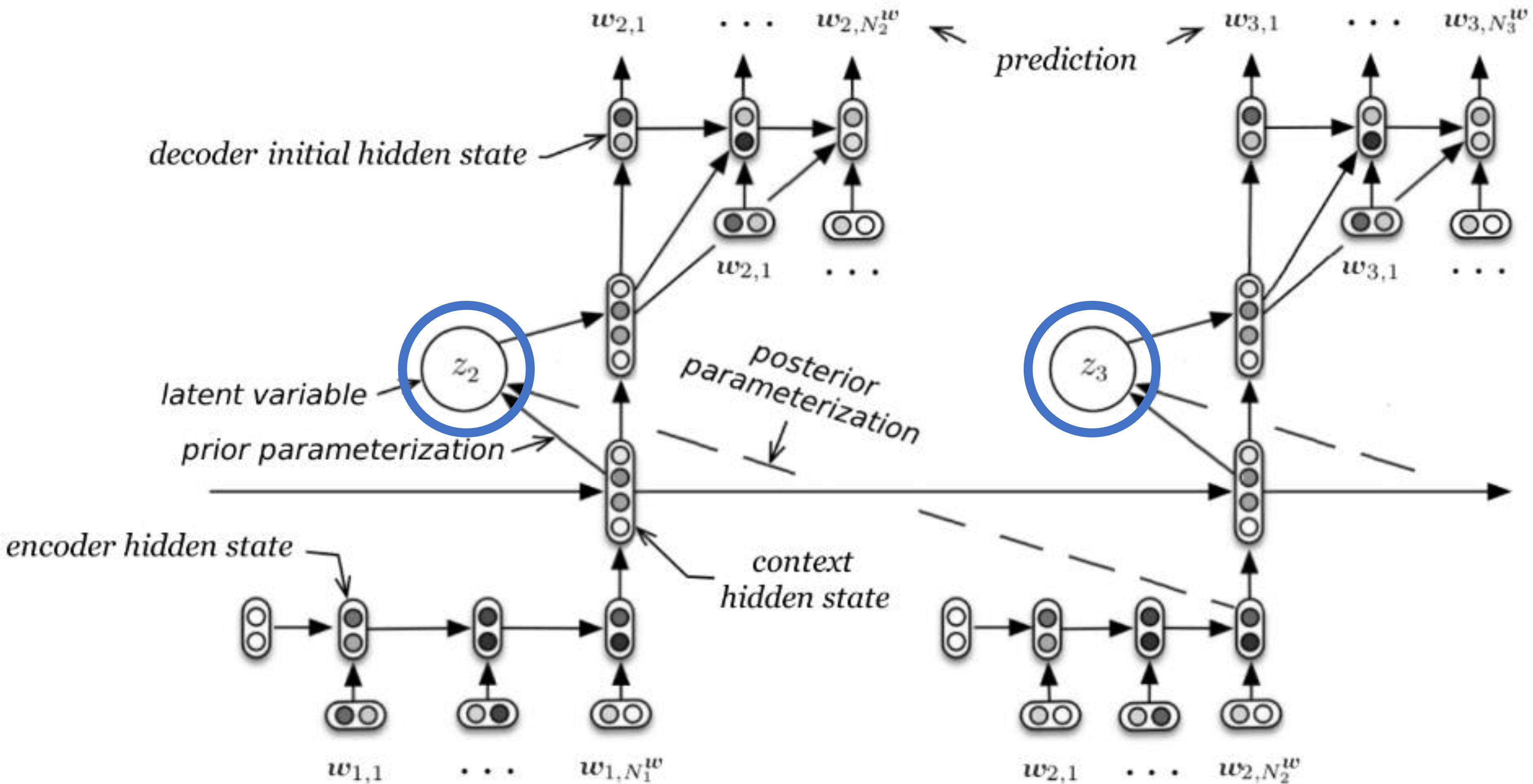
- **Hierarchical Encoder-Decoder (HRED)** [Serban+ 16]



Encodes:
utterance (word by word) +
conversation (turn by turn)

Challenge: Long conversational context

- **Hierarchical** **Latent Variable** **Encoder-Decoder (VHRED)** [Serban+ 17]
  - Adds a latent variable to the decoder
  - Trained by maximizing variational lower-bound on the log-likelihood



Related to persona model [Li+ 2016b]:

Deals with 1-N problem, but unsupervisedly.

Hierarchical Encoders and Decoders: Evaluation

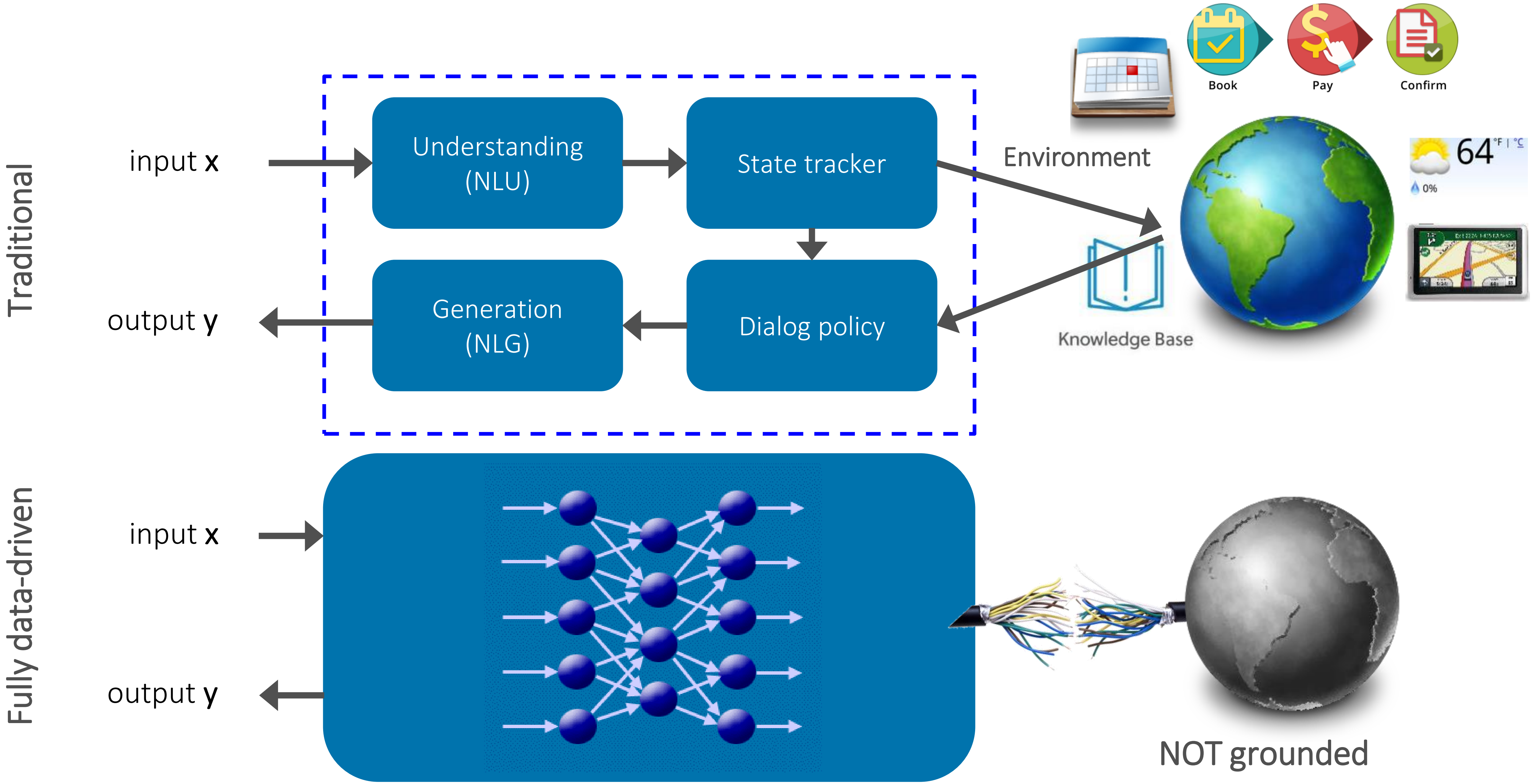
[Serban+ 17]

Opponent	Wins	Losses	Ties
Short Contexts			
VHRED vs LSTM	32.3 ±2.4	**42.5 ±2.6***	25.2 ±2.3
VHRED vs HRED	**42.0 ±2.8***	31.9 ±2.6	26.2 ±2.5
VHRED vs TF-IDF	**51.6 ±3.3***	17.9 ±2.5	30.4 ±3.0
Long Contexts			
VHRED vs LSTM	**41.9 ±2.2***	36.8 ±2.2	21.3 ±1.9
VHRED vs HRED	**41.5 ±2.8***	29.4 ±2.6	29.1 ±2.6
VHRED vs TF-IDF	**47.9 ±3.4***	11.7 ±2.2	40.3 ±3.4

Outline

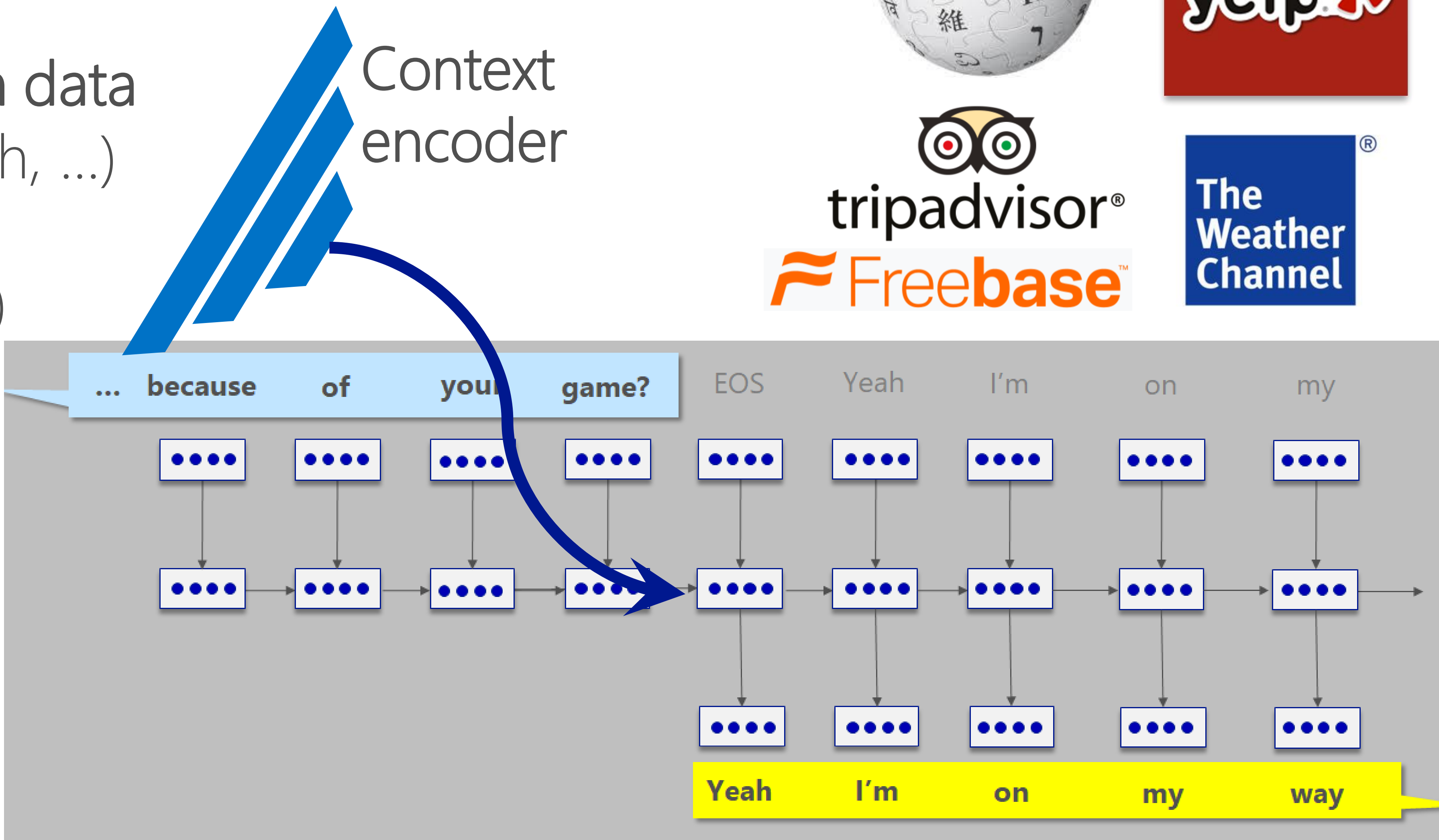
- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogue
- **Part 4: Fully data-driven conversation models and chatbots**
  - **E2E neural conversation models**
  - **Challenges and remedies**
  - **Grounded conversation models**
  - **Beyond supervised learning**
  - **Data and evaluation**
  - **Chatbots in public**
  - **Future work**

Towards Grounded E2E Conversation Models



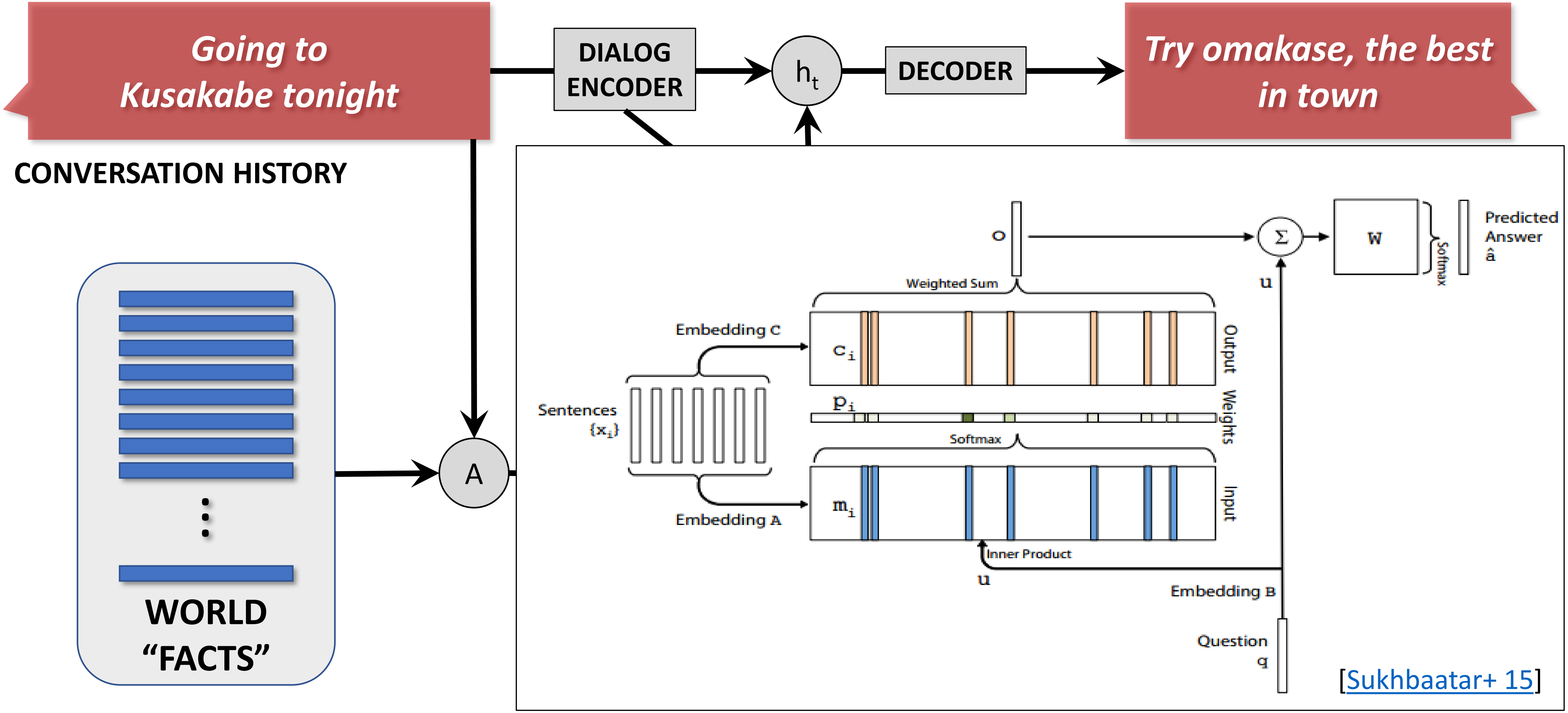
E2E Conversation Models in the real world

Personalization data
(ID, social graph, ...)
Device sensors
(GPS, vision, ...)
External
"knowledge"



Knowledge-Grounded Neural Conversation Model

[Ghazvininejad+ 17; Agarwal+ 18; Liu+ 18]



Sample knowledge-grounded responses

I would give ALMOST anything for some [*Mexican restaurant*] right now.

Me too. Jalapeno sauce is really good.

Obsessed with [*jewelry company*] :-*

oh my gosh obsessed with their bracelets and the meaning behind them!

Visiting the celebs at Los Angeles airport - [...] w/ 70 others

Nice airport terminal. Have a safe flight.

<u>Experimental results (23M conversations):</u> outperforms competitive neural baseline (human + automatic eval)

Conversations grounded in *Full-Length* Documents

Vesna Vulović

From Wikipedia, the free encyclopedia

Vesna Vulović (Serbian Cyrillic: Весна Вуловић; pronounced [ˈʋeːsna ˈʋuːlɔʋitɕ]; 3 January 1950 – 23 December 2016) was a Serbian flight attendant. She holds the Guinness world record for surviving the highest fall without a parachute: 10,160 metres (33,330 ft). Her fall took place after an explosion tore through the baggage compartment of JAT Flight 367 on 26 January 1972, causing it to crash near Srbská Kamenice, Czechoslovakia. She was the sole survivor of the crash that air safety investigators attributed to a briefcase bomb. The Yugoslav authorities suspected that Croatian nationalists were to blame, but no one was ever arrested.

Following the crash, Vulović spent days in a coma and was hospitalized for several months. She suffered a fractured skull, three broken vertebrae, two broken legs, broken ribs and a fractured pelvis. These injuries resulted in her being temporarily paralyzed from the waist down. She made an almost complete recovery but continued to walk with a limp. Vulović maintained that she had no memory of the incident and thus had no qualms about flying in the

Vesna Vulović

Vulović in the early 1970s

Born	3 January 1950 Belgrade, PR Serbia, FPR Yugoslavia
Died	23 December 2016 (aged 66) Belgrade, Serbia

A woman fell 30,000 feet from an airplane and survived.

The page states that a 2009 report found the plane only fell several hundred meters.

Well if she only fell a few hundred meters and survived then I 'm not impressed at all.

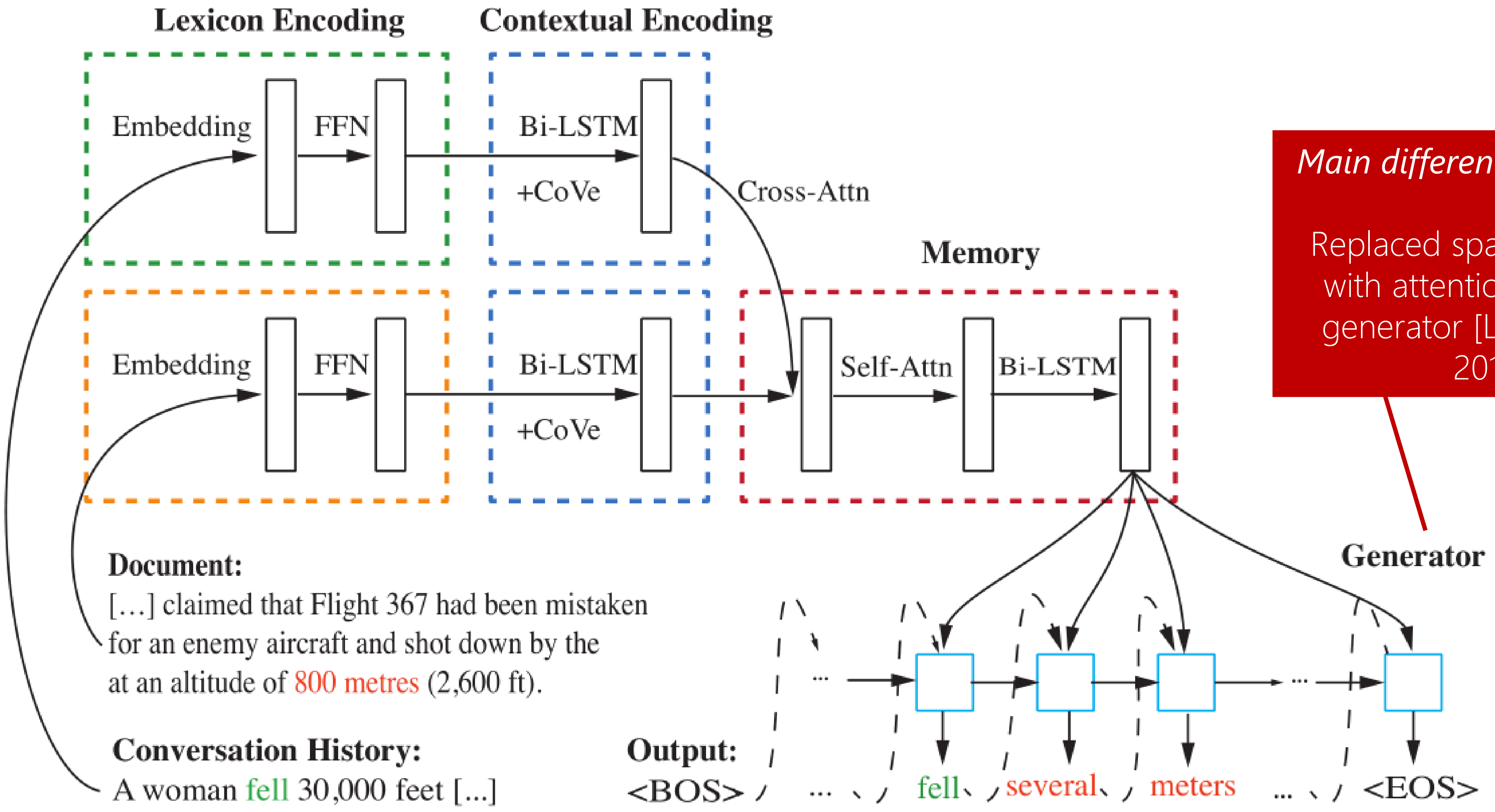
Few hundred meters is still pretty incredible , but quite a bit different than 10,000 meters.

Task: Generate a **human-like response** that is not only conversationally appropriate, but also **informative** (→ useful task) and **grounded** (-> evaluation closer to MRC).

[Dinan+ 19; Qin+ 19]

Models with Document-Level Grounding

Machine Reading Comprehension-based Model [Qin+ 19]:



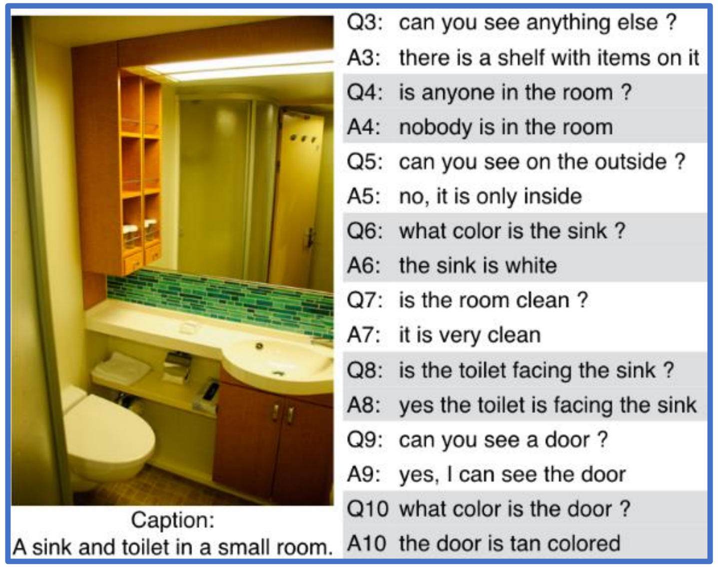
Main difference with MRC:

Replaced span prediction with attention recurrent generator [Luong et al., 2015]

Grounded E2E Dialogue Systems

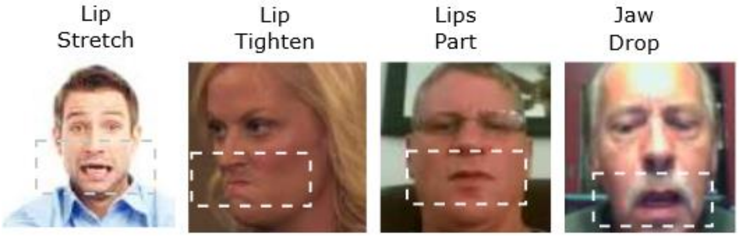
- **Grounding: images**

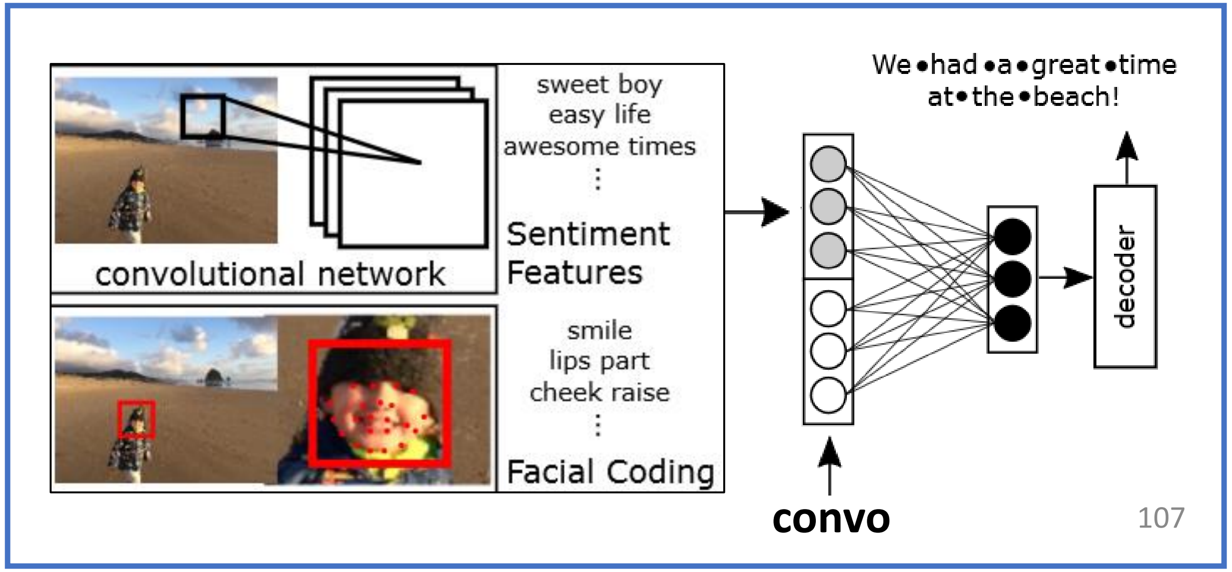
  Conversations around images e.g.,
  Q-As [Das+ 16] or chat [Mostafazadeh+ 17]



- **Grounding: affect** [Huber+ 18]

  facial actions influence response





Beyond supervised learning

- **Limitations of SL for E2E dialogue:**
  - Train on human-human data, test with human-machine (Twitter-ese often not what we want at test time.)
  - Optimizes for immediate reward $p(T_N \mid \dots T_{N-1})$, not long-term reward
  - No user feedback loop

- **Emergence of reinforcement learning (RL) for E2E dialogue**
  - Tries to promote long-term dialogue success

Deep Reinforcement Learning for E2E Dialogue

[Li+ 16c]

- REINFORCE algorithm [Williams+ 92]

$$J(\theta) = \mathbb{E}[R(s_1, s_2, ..., s_N)]$$

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, ..., s_N) \boxed{R(s_1, s_2, ..., s_N)}$$

reward function

$$\nabla J(\theta) = \nabla \log \prod_i \boxed{p(s_i | s_{i-1})} R(s_1, s_2, ..., s_N)$$

what we want to learn

- Reward functions:
  1. Ease of answering: $-Pr(\text{Dull Response} | s_i)$
  2. Information flow: $-\log \text{Sigmoid} \cos(s_1, s_2)$
  3. Meaningfulness: $\log p_{\text{seq2seq}}(s_1 | s_2) + \log p_{\text{seq2seq}}(s_2 | s_1)$

Simulation (without RL)

See you later!

See you later!

See you later!

See you later!

See you later!

See you later!

See you later!

110

Simulation (with RL)

How old are you ?

i 'm 4, why are you asking ?

I thought you were 12 .

What made you think so ?

I don’t know what you are talking about .

You don’t know what you are saying.

I don’t know what you are talking about .

Deep RL: Evaluation

- MTurk evaluation (500 responses)

Setting	RL-win	RL-lose	Tie
single-turn general quality	0.40	0.36	0.24
single-turn ease to answer	0.52	0.23	0.25
multi-turn general quality	0.72	0.12	0.16

Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogue
- **Part 4: Fully data-driven conversation models and chatbots**
  - **E2E neural conversation models**
  - **Challenges and remedies**
  - **Grounded conversation models**
  - **Beyond supervised learning**
  - **Data and evaluation**
  - **Chatbots in public**
  - **Future work**

Conversational datasets
(for social bots, E2E dialogue research)

- Survey on dialogue datasets [Serban+ 15]

Name	Type / Topics	Size
Reddit	Unrestricted	3.2B dialog turns (growing)
Twitter	Unrestricted	N/A (growing)
OpenSubtitles	Movie subtitles	1B words
Ubuntu Dialogue Corpus	Chat on Ubuntu OS	100M words
Ubuntu Chat Corpus	Chat on Ubuntu OS	2B words
Persona-Chat Corpus	Crowdsourced / personalized	164k dialog turns

Evaluating E2E Dialogue Systems

- **Human evaluation (crowdsourcing):**

Context: ... *Because of your game?*
Input: *Yeah, I'm on my way now*
Response: **Ok good luck!**

Is **this** a good[1] response?

Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree
○	○	○	◉	○

[1]: replaced as appropriate (relevant, interesting,…)

- **Automatic evaluation:**
  Less expensive, but is it reliable?

Machine-Translation-Based Metrics

- **BLEU** [Papineni+ 02]: ngram overlap metric

Reference:	John	resigned	yesterday	.
System:	Yesterday ,	John	quit .	

$$\text{BLEU} = \text{BP} \cdot \exp\left(\sum_n \log p_n\right)$$

$$p_n = \frac{\sum_i \sum_{g \in n\text{-grams}(h_i)} \max_j \{\#_g(h_i, r_{i,j})\}}{\sum_i \sum_{g \in n\text{-grams}(h_i)} \#_g(h_i)}$$

- **NIST** [Doddington+ 02]
  - Seldom used in dialogue, but copes with **blandness issue**
  - Considers info gain of each ngram: score(*interesting calculation*) >> score(*of the*)
- **METEOR**
  - Accounts for synonyms, paraphrases, etc.

The challenge with MT-based metrics

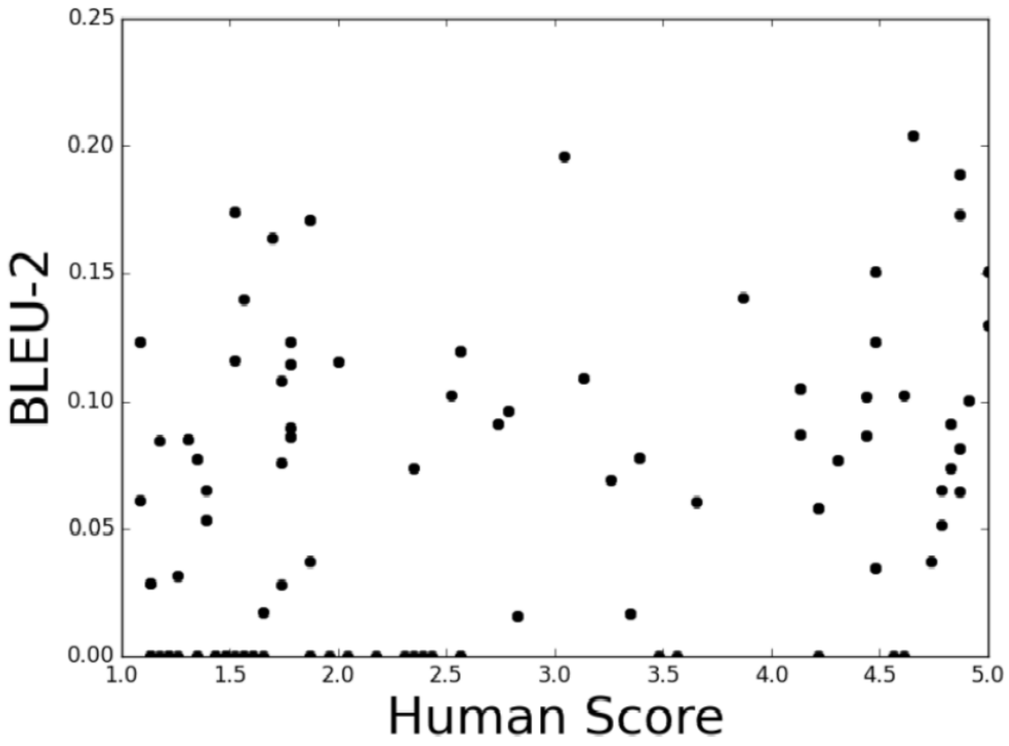
Input:	*How are you?*
Response (gold):	**I 'm good , thanks .**

Response A:	**Good thanks** !	***Semantically*** *equivalent (as in Machine Translation)*
Response B:	Doing pretty **good thanks**	
Response C:	Doing well thank you !	
Response D:	Fantastic **.** How are you ?	***Pragmatically*** *appropriate*
Response E:	**I 'm** getting sick again **.**	
Response F:	Bored **.** you ?	
Response G:	Sleepy **.**	
Response H:	Terrible tbh	

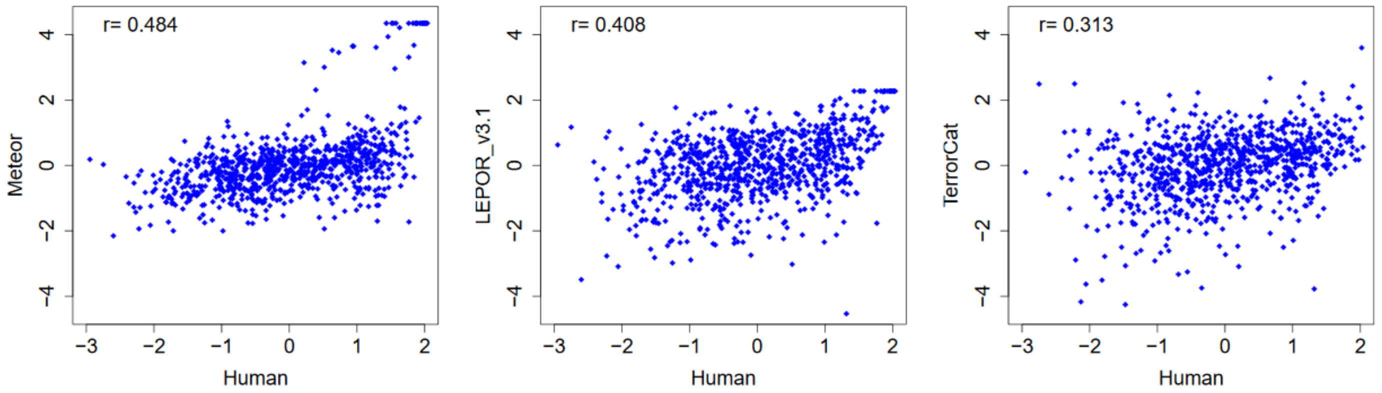
Many false negative!

Sentence-level correlation of MT metrics

- Poor correlation with human judgments:



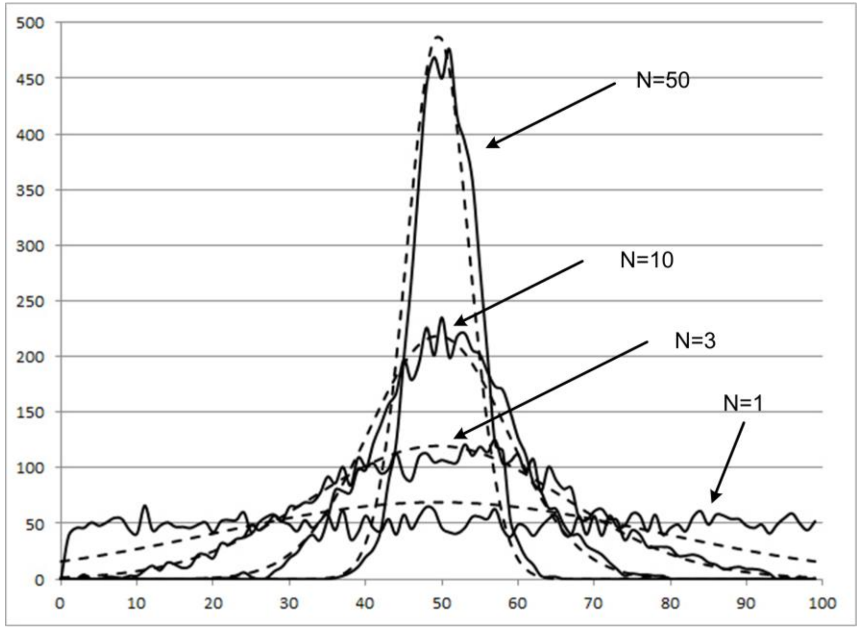
Dialogue task
“How NOT to evaluate dialogue systems”
[Liu+ 16]



But same problem even
for **Translation task**
[Graham +15]

The importance of sample size

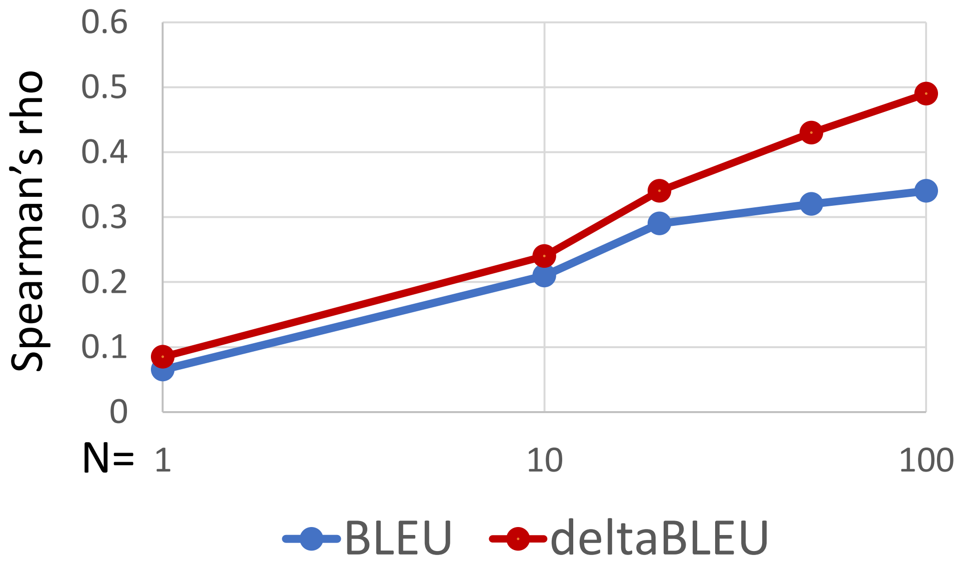
- **MT metrics were NOT designed to operate at the sentence level:**
  - BLEU [Papineni+ 02] == "corpus-level BLEU"
  - Statistical Significant Tests for MT [Koehn 06; etc.]: BLEU not reliable with sample size < 600, *even* for Machine Translation (easier task)
- **Central Limit Theorem (CLT) argument:**
  - Matching against reference (e.g., n-grams) is brittle → greater variance
  - Remedy: reduce variance by increasing sample size (CLT), i.e., **corpus-level BLEU**



(Figure from [Brooks+ 12])

Corpus-level Correlation

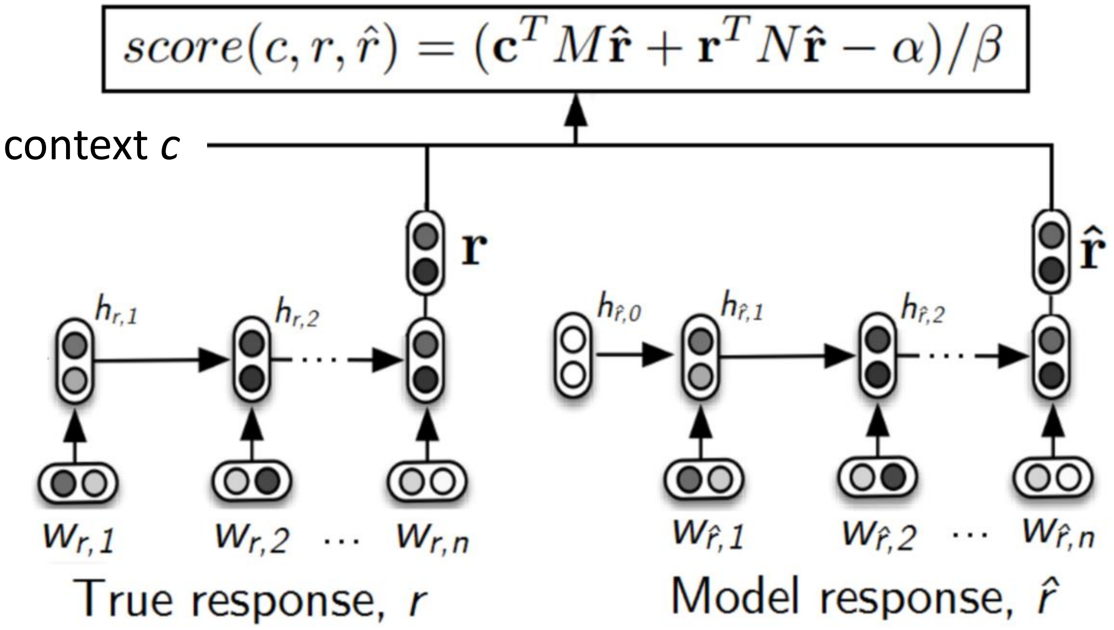
- Generally good for **Machine Translation** (MT):
  - Spearman's rho of 0.8 to 0.9 for BLEU, METEOR [Przybocki+ 08]
- Can it work for **Dialogue**?
  - Currently no definite answer, as corpus-level human judgments very expensive.
  - Experiments with smaller N [Galley+ 15]:



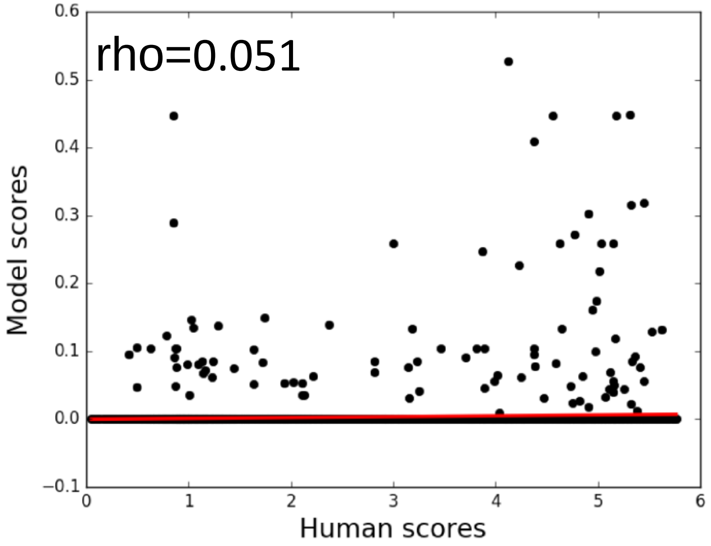
deltaBLEU = human-rating weighted version of BLEU [Galley+ 15]

Trainable Metric

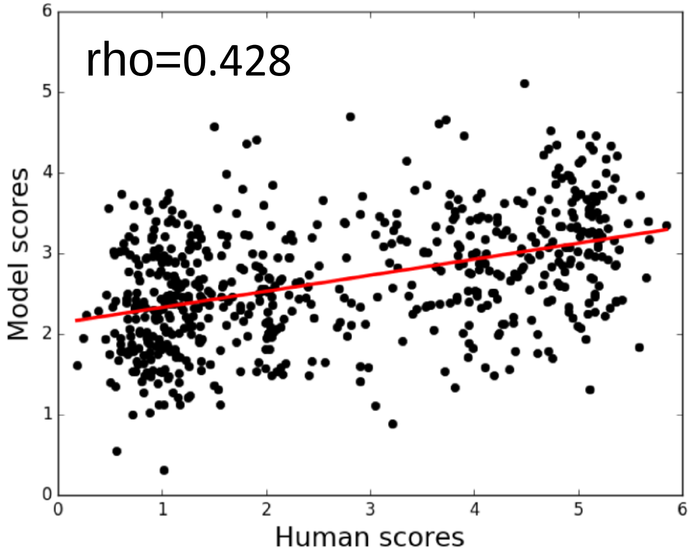
- **Towards an automatic turning test** [Lowe+ 17]: ADEM: Metric based on hierarchical RNN (VHRED)



BLEU-2 (N=1)

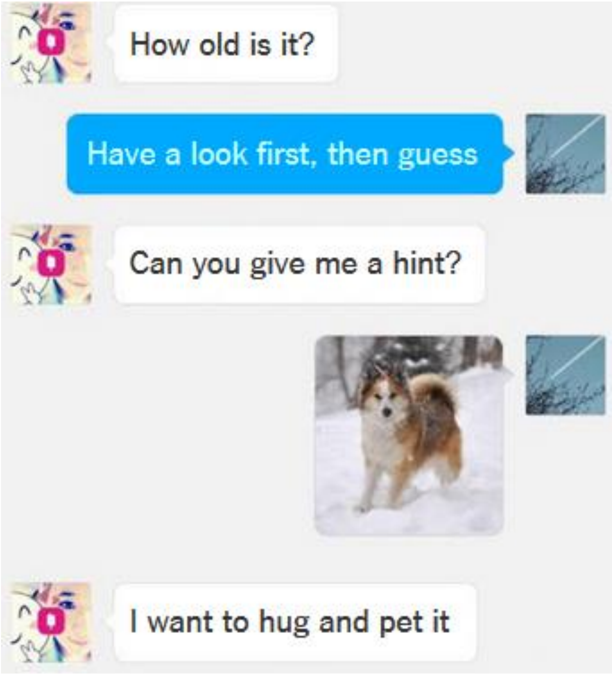


ADEM (N=1)

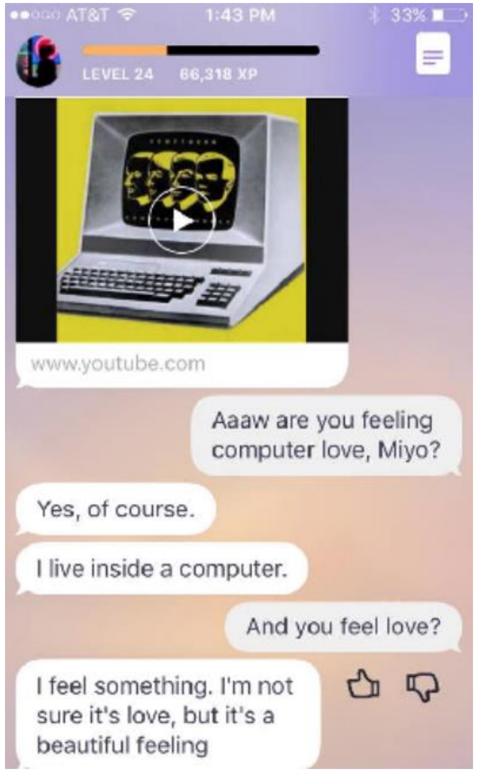


Social Bots: commercial systems

- **For end users:**
  - Amazon Alexa
    (trigger: say “Alexa, let’s chat”)
  - Microsoft XiaoIce [Zhou+ 2018]
  - Microsoft Zo (on Kik)
  - Replika.ai [system description]



XiaoIce
(translated from Chinese)



Replika.ai

For bot developers:

- Microsoft Personality Chat (includes speaker embedding LSTM)

Microsoft

Cognitive Services Labs Home Keys Documentation Stack Overflow Feedback

Query:
I am going on stage :)

Personality: Professional — Clear Chat

Top Responses & Ranked Score

What are you going to perform?	0.71
What are you going to sing?	0.69
Good luck with that.	0.67
I hope you have a great time.	0.67
Good luck with that...	0.65

Personality: Friendly — Clear Chat

Top Responses & Ranked Score

Oh my gosh that will be amazing	0.69
Oh my gosh that's amazing. Have a great time	0.65
I'm sure you will have a great time	0.65
Oh my gosh that will be amazing. Are you going	0.65
Oh my gosh that's brilliant. Have a great time	0.65

Personality: Humorous — Clear Chat

Top Responses & Ranked Score

With who?! D:	0.77
Really?! Oh my gosh I'm so jealous!	0.77
Oh my gosh you're so lucky! D:	0.77
That's awesome! Good luck! C:	0.77
Good luck! I'm sure you'll be amazing!	0.75

https://labs.cognitive.microsoft.com/en-us/project-personality-chat

Open Benchmarks

- **Alexa Challenge** (2017-)
  - Academic competition, 15 sponsored teams in 2017, 8 in 2018
  - $250,000 research grant (2018)
  - Proceedings [2017, 2018]

- **Dialogue System Technology Challenge (DSTC)** (2013-)
  (formerly Dialogue State Tracking Challenge)
  Focused this year on **grounded conversation**:
  Visual-Scene [Hori +18], knowledge grounding [Galley +18]

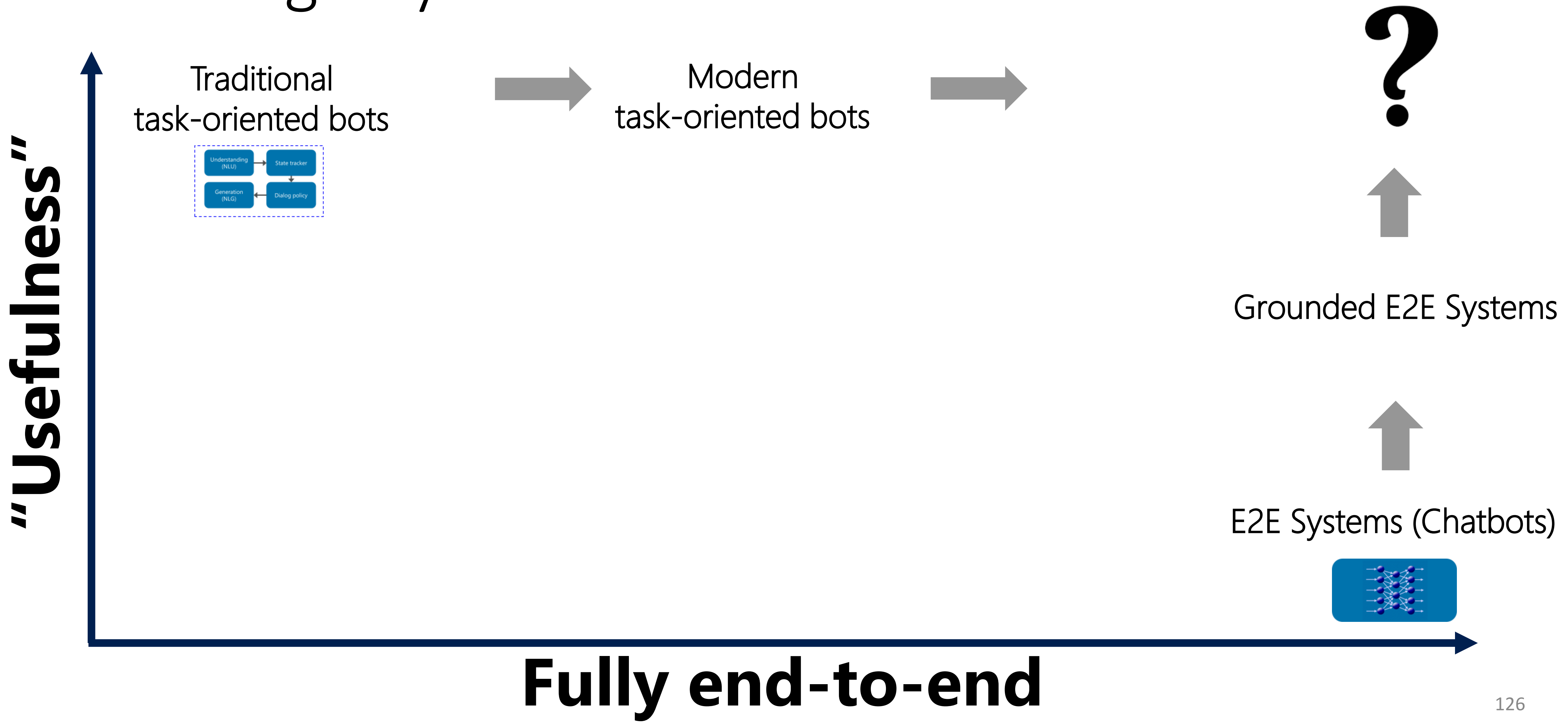
- **Conversational Intelligence Challenge (ConvAI)** (2017-)
  Last occurrence focused on **personalized chat** (Persona-Chat dataset)

Conclusions

- E2E Neural Conversation Models
  - Learn the backbone or shell of open-domain natural conversation
  - Face significant challenges (blandness, consistency, long context), but alleviated using better models and objectives (e.g., MMI and HRED)

- Grounded conversational AI models
  - Exploit external textual knowledge, device sensors (e.g., images), personal information
  - Produce more informational and "**useful**" dialogues

Moving beyond chitchat



Fully Data-driven Response Generation:

Challenges and future work

Better objective functions and evaluation metrics

- **Lack of good objective or reward functions is a challenge for SL and RL:**
  - **MLE** causes blandness (mitigated by MMI)
  - **Evaluation metrics** (BLEU, METEOR, etc.) reliable only on large datasets
    → **expensive for optimization** (e.g., sequence-level training [Ranzato+ 15])
  - **RL reward** functions currently too **ad-hoc**

- **Final system evaluation:**
  - Still need human evaluation
  - Corpus-level metrics (BLEU, METEOR, etc.): How effective are they really?

Better leverage **heterogeneous** data

English sentence 1 → French sentence 1

⋮

English sentence N → French sentence N

most NLP / AI problems
(homogeneous data)

reddit

query 1 → response 1

⋮

query N → response N

general domain dialog

WIKIPEDIA IMDb

FOURSQUARE tripadvisor®

much of world knowledge in non-conversational form (often unstructured)

The Service Line · Microsoft Teams · slack · Ask a Librarian · TECH SUPPORT

in-domain data (e.g., decision making, task-oriented)

conversational AI
(heterogeneous data)

Thank you

Contact Information:

Jianfeng Gao http://research.microsoft.com/~jfgao

Michel Galley http://research.microsoft.com/~mgalley

Slides:
https://icml.cc/Conferences/2019/Schedule

Journal paper version of this tutorial:
https://www.nowpublishers.com/article/Details/INR-074 (final)

https://arxiv.org/abs/1809.08267 (preprint)