



Building a Conversational Agent Overnight with Dialogue Self-Play

Pararth Shah

Research Engineer, Google AI

Stanford MS CS '15

February 2018

Outline

- **Problem:** Building richer conversational interactions
- **Key idea:** Dialogue as a collaborative game
- **Approach:** Combining automation & human intelligence
- **Evaluation & Applications**

Complexity in conversational agents

Task Complexity

Multi-domain

Transactional

Informational



Language Complexity

Implicature

Entailment

Coreference

Basic

Web search

Question Answering

Long-form dialogue

Discourse Complexity

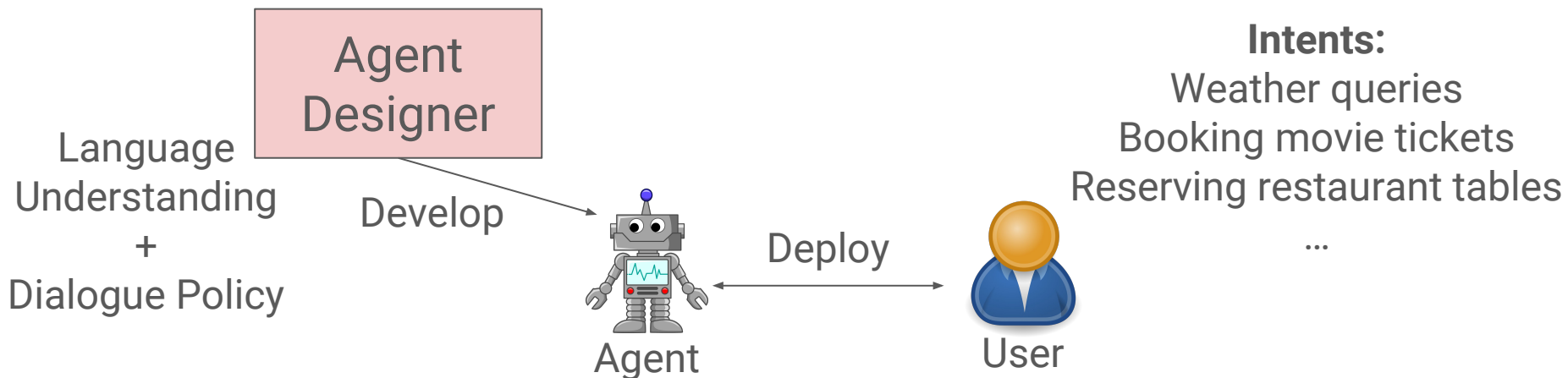
“Tell me the weather”

“How many trophies did he win?”

“I want to get lunch with her. Book us a table.”

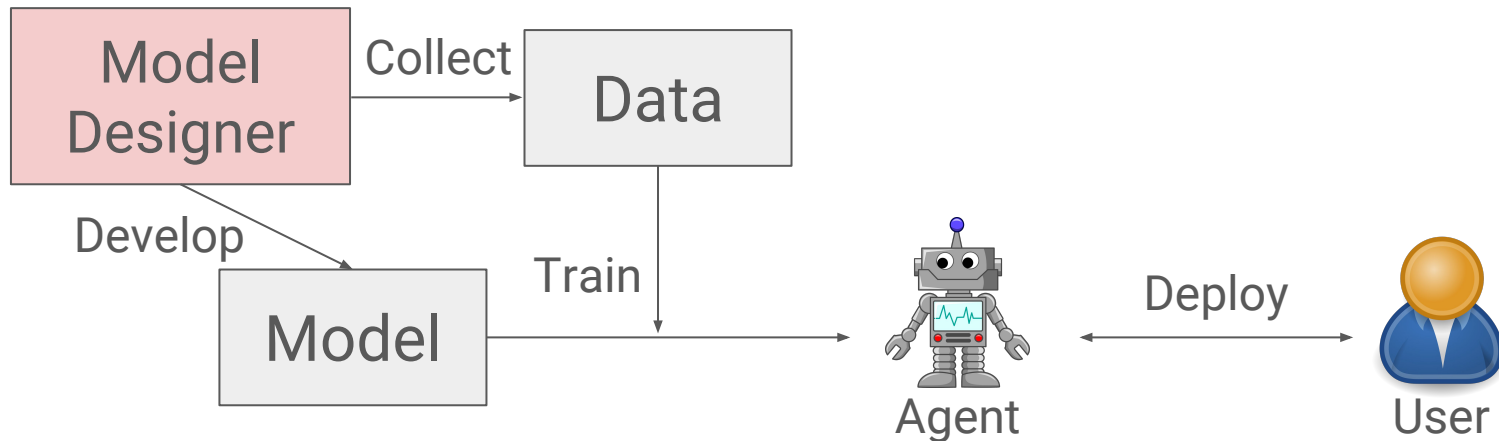
“Explain the modern scientific method’s roots in Aristotle’s philosophy”

Industry practice: Engineer each new capability



- + Full control over agent behavior
- + Guaranteed coverage of critical interactions
- Low recall in unanticipated interactions
- Agent does not learn from mistakes

Research focus: Models trained from data



- + Learn interactions from users/crowd
- + Flexible agent, can improve with more data
- Dataset collection and annotation expensive
- Little control over agent behavior

How to achieve both control & flexibility at scale?

Control

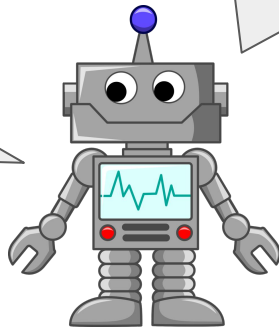
Guaranteed coverage of critical user journeys

Discourse complexity

Improve language understanding from data and experience

Flexibility

Language complexity



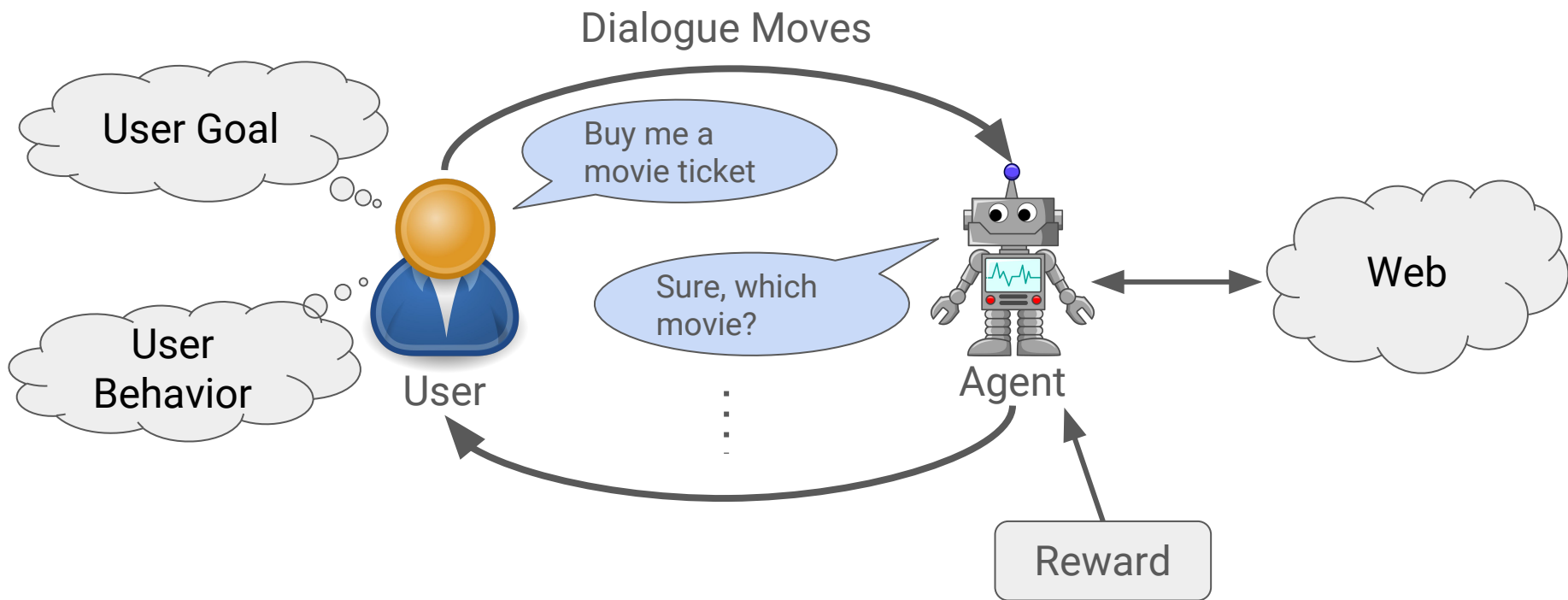
Agent

Scale

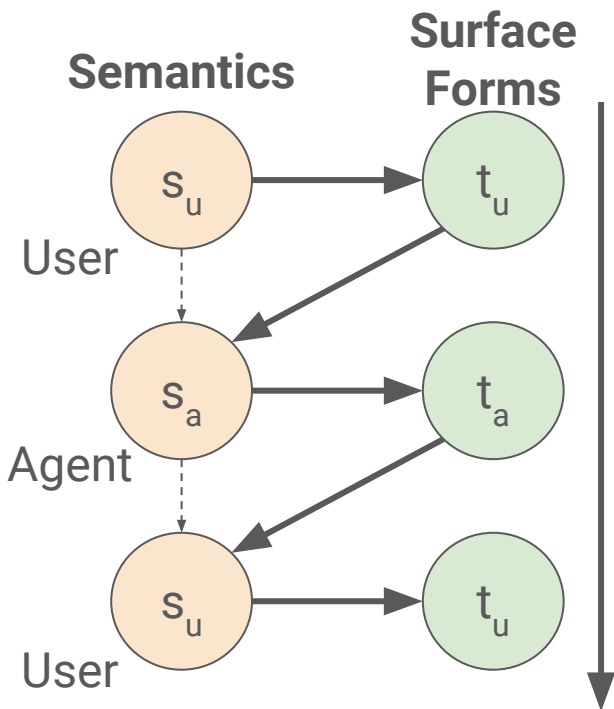
Efficiently add new capabilities

Task complexity

Idea 1: Dialogue is a collaborative game



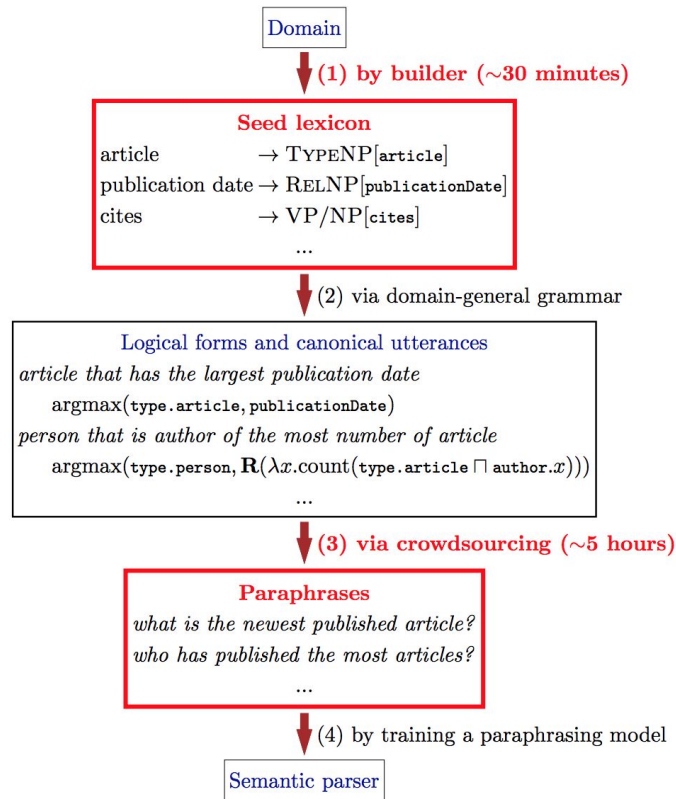
Idea 2: Separate task/discourse from language



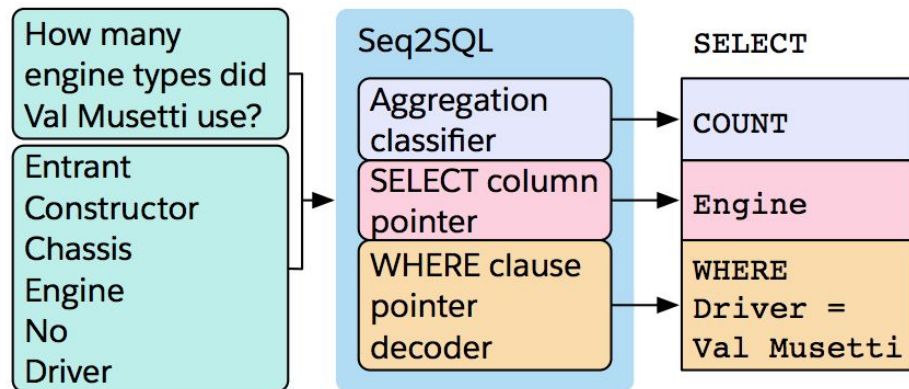
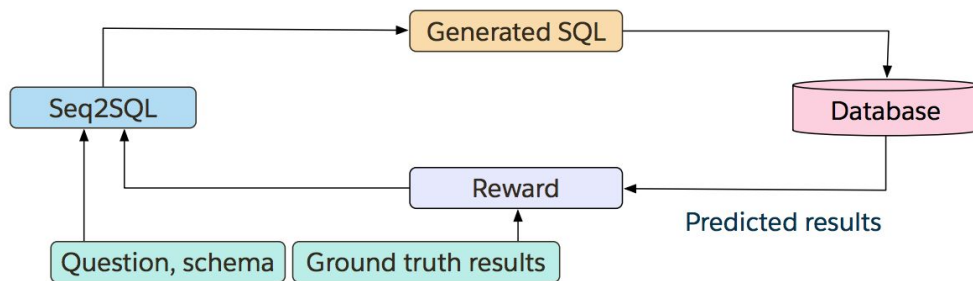
“Dialogue Outline”

Semantic annotations (s)	Template utterances (t)
A: greeting()	Greeting.
U: inform(intent=book_movie, name=Inside Out, date=tomorrow, num_tickets=2)	Book movie with name is Inside Out and date is tomorrow and num tickets is 2.
A: ack() request(time)	OK. Provide time.
U: inform(time=evening)	Time is evening.
...	...

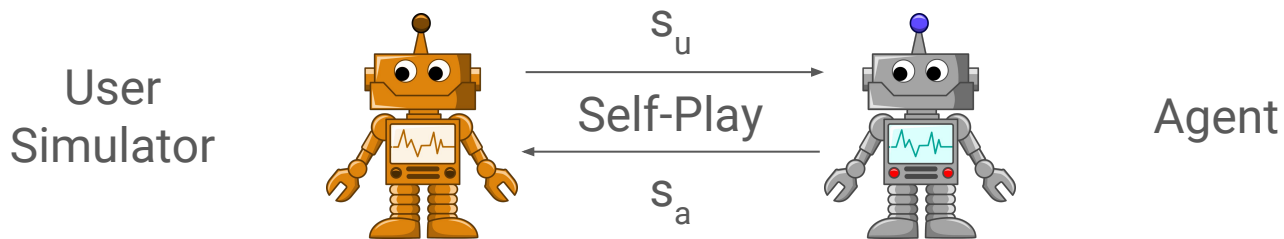
Yushi Wang, Jonathan Berant, and Percy Liang. "Building a semantic parser overnight." In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics. 2015.



Victor Zhong, Caiming Xiong, and Richard Socher. "Seq2SQL: Generating Structured Queries from Natural Language using Reinforcement Learning." arXiv preprint arXiv:1709.00103 (2017).

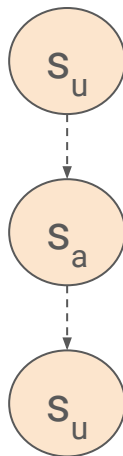


Idea 3: User simulation & dialogue self-play



Generative model of user's dialogue actions at semantic level

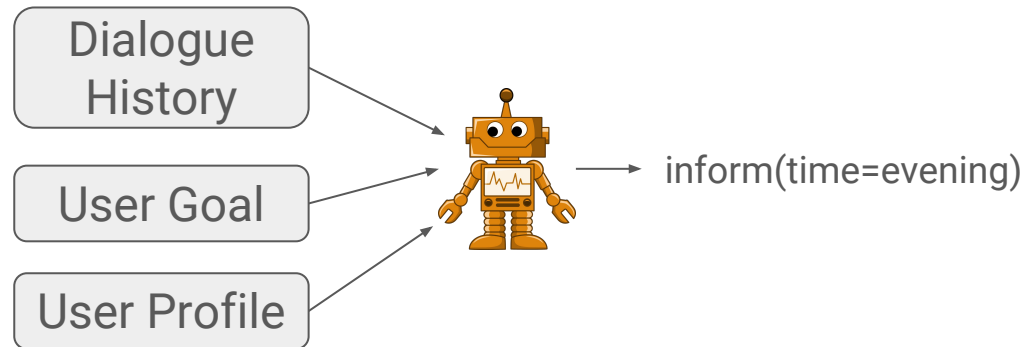
Parameterized by user goal and user profile



Iteratively sample dialogue moves from user and system agent

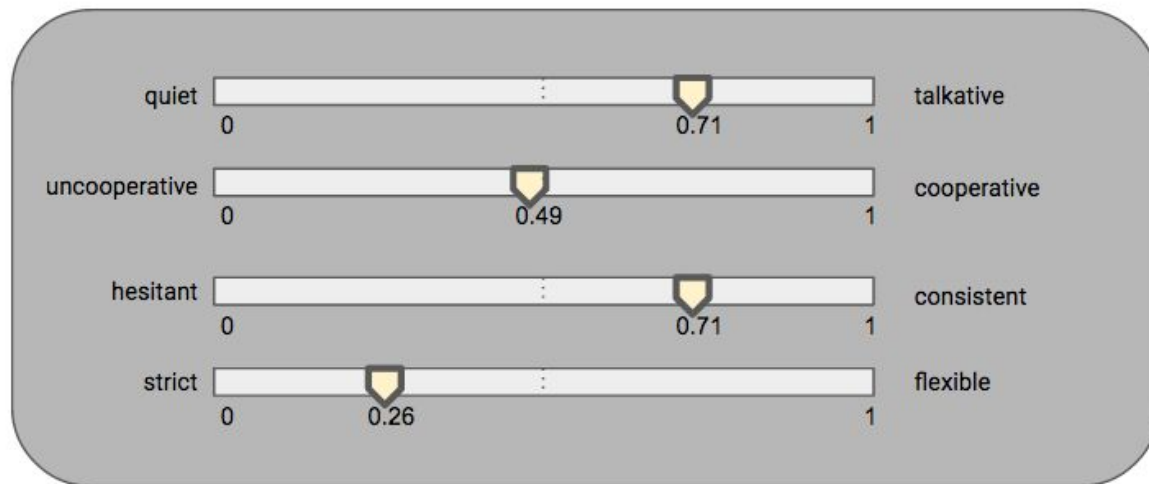
Exhaustively explore space of outlines for the task

User simulator



User profile parameters:

- Talkativeness
- Cooperativeness
- Consistency
- Flexibility



Idea 4: Crowd-sourced dialogue rewrites

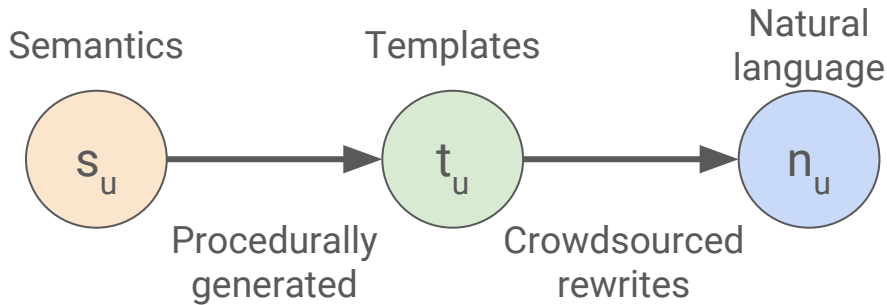
Instructions

You'll be shown a **very unnatural computer generated conversation** between a user and an assistant.

Your task is to paraphrase the messages in the original conversation in order to create a new conversation that has the same meaning as the original conversation between a user and a professional assistant.

Feel free to be creative with your paraphrased messages, as long as they meet the following **Important requirements**:

- User messages need to look like something you would type in a chat window.
- Assistant messages need to look formal – just like chat replies from a human customer service agent.
- **All of your paraphrased messages must have the same meaning as the original messages.**
- Your paraphrased conversation needs to make sense on it's own.



Task

t_u	n_u
Assistant: greetings	Hello
User: greetings AND buy movie tickets AND date is Wednesday	I would like to buy movie tickets for Wednesday.
Assistant: provide preference for: theatre name and time and number of tickets and movie	What theatre, movie, time, and number of tickets would you like?
User: unable to understand, must re-phrase your message	Sorry, I didn't understand. Could you repeat that?

Outline		Paraphrase
Semantic annotations	Template utterances	Natural utterances
A: greeting()	Greeting.	Hi, how can I help you?
U: inform(intent=book_movie, name=Inside Out, date=tomorrow, num_tickets=2)	Book movie with name is Inside Out and date is tomorrow and num tickets is 2.	I want to buy two tickets for Inside Out for tomorrow.
A: ack() request(time)	OK. Provide time.	Alright. What time would you like to see the movie?
U: inform(time=evening)	Time is evening.	Anytime during the evening works for me.
A: offer(theatre=Cinemark 16, time=6pm)	Offer theatre is Cinemark 16 and time is 6pm.	How about the 6pm show at Cinemark 16?
U: affirm() inform(intent=find_rest, meal=dinner, location=near theatre)	Agree. Find restaurant with meal is dinner and location is near theatre.	That sounds good. Also, I want to get dinner near that theatre.

Combining automation & human intelligence

Machines Talking to Machines (M2M)

Task Schema:

movies: name, theatre, date, time, ...
restaurants: name, cuisine, price, location, ...

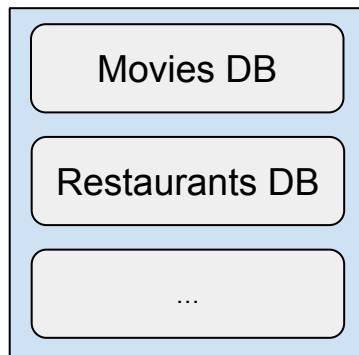
User Profile:

```
verbose: 0.8  
flexible: 0.5  
asks_for_repeat: 0.3  
...
```

User Goal:

```
book_movie:  
name=Inside Out  
theatre=dont_care  
time=evening  
...  
reserve_restaurant:  
location=near theatre  
time=after the movie  
...
```

API Client:

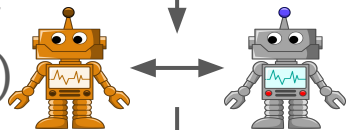


(a) Task Specification



Scenarios

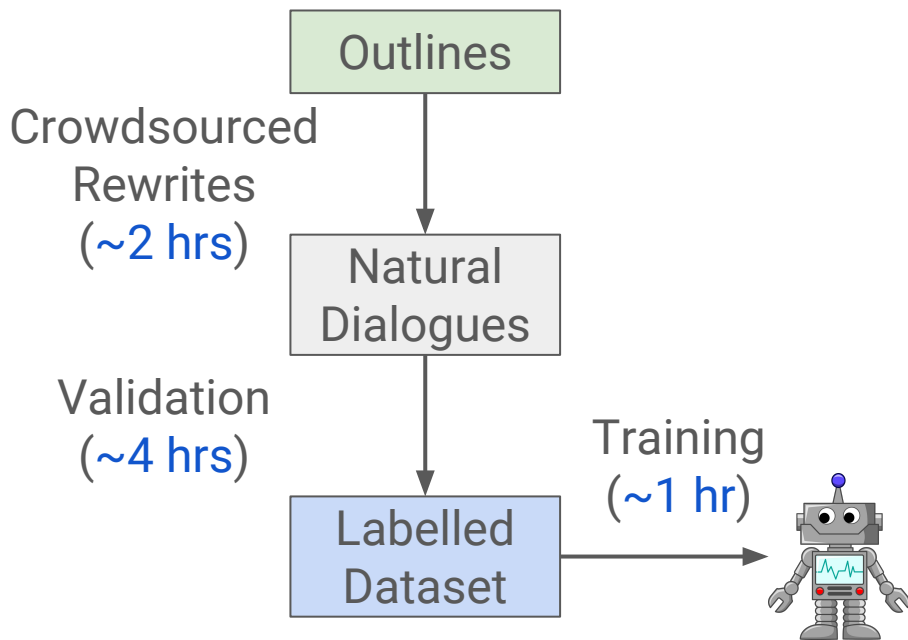
Self-Play
(~5 mins)



Outlines

Combining automation & human intelligence

Machines Talking to Machines (M2M)



Automation: Dialogue self-play (**5mins**)

- Explore dialogue outlines
- Discourse complexity

Human IQ: Crowdsourcing (**6hrs**)

- Add natural sounding utterances
- Language complexity

Scale to new tasks **overnight** (**8hrs**)

- With task schema and API client
- Task complexity

Dataset release

github.com/google-research-datasets/simulated-dialogue

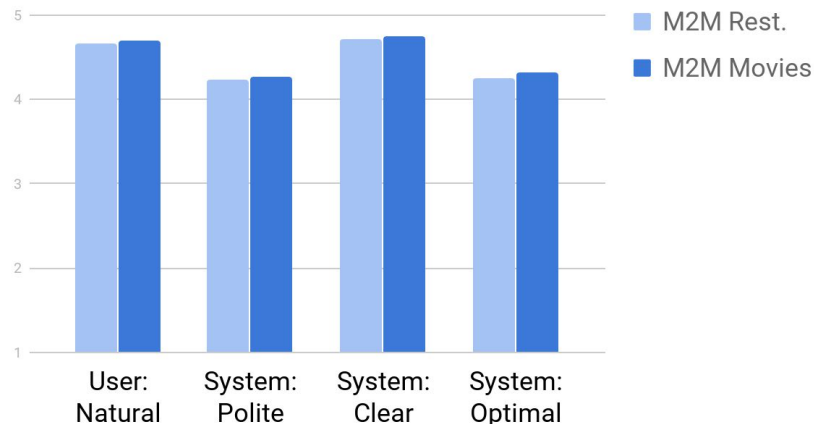
Dataset	Schema	Train	Dev	Test
M2M Restaurants	price_range, location, restaurant_name, category, num_people, date, time	1116	349	775
M2M Movies	theatre_name, movie, date, time, num_people	384	120	264

Pararth Shah, Dilek Hakkani-Tür, Gokhan Tur, Abhinav Rastogi, Ankur Bapna, Neha Nayak, and Larry Heck. "Building a Conversational Agent Overnight with Dialogue Self-Play." arXiv preprint arXiv:1801.04871 (2018).

Data quality evaluation

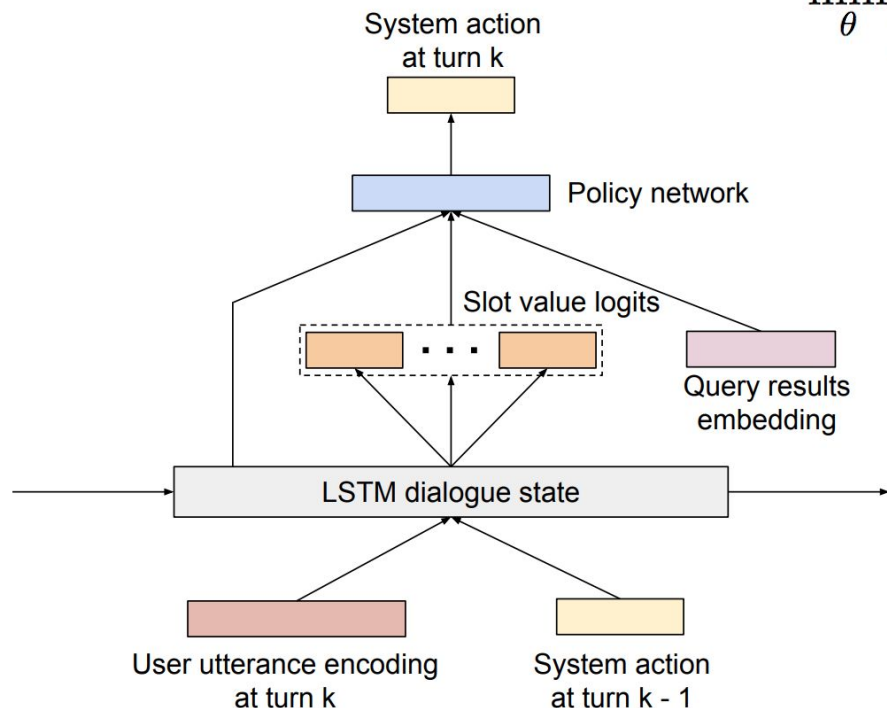
Metric	DSTC 2 (Train)	M2M Rest. (Train)
Unique tokens / Total tokens	0.0049	0.0092
Unique bigrams / Total tokens	0.0177	0.0670
Unique transitions / Total turns	0.0982	0.2646
Unique sub-dialogues (k=3) / Total sub-dialogues (k=3)	0.1831	0.3145
Unique sub-dialogues (k=5) / Total sub-dialogues (k=5)	0.5621	0.7061

Human evaluation



Pararth Shah, Dilek Hakkani-Tür, Gokhan Tur, Abhinav Rastogi, Ankur Bapna, Neha Nayak, and Larry Heck. "Building a Conversational Agent Overnight with Dialogue Self-Play." arXiv preprint arXiv:1801.04871 (2018).

Agent architecture



Goal slot values

$$\min_{\theta} \sum_{k=1}^K - \left[\sum_{m=1}^M \lambda_{l^m} \log P(l_k^{m*} | \mathbf{U}_{\leq k}, \mathbf{A}_{<k}, \mathbf{E}_{<k}; \theta) + \lambda_a \log P(a_k^* | \mathbf{U}_{\leq k}, \mathbf{A}_{<k}, \mathbf{E}_{\leq k}; \theta) \right]$$

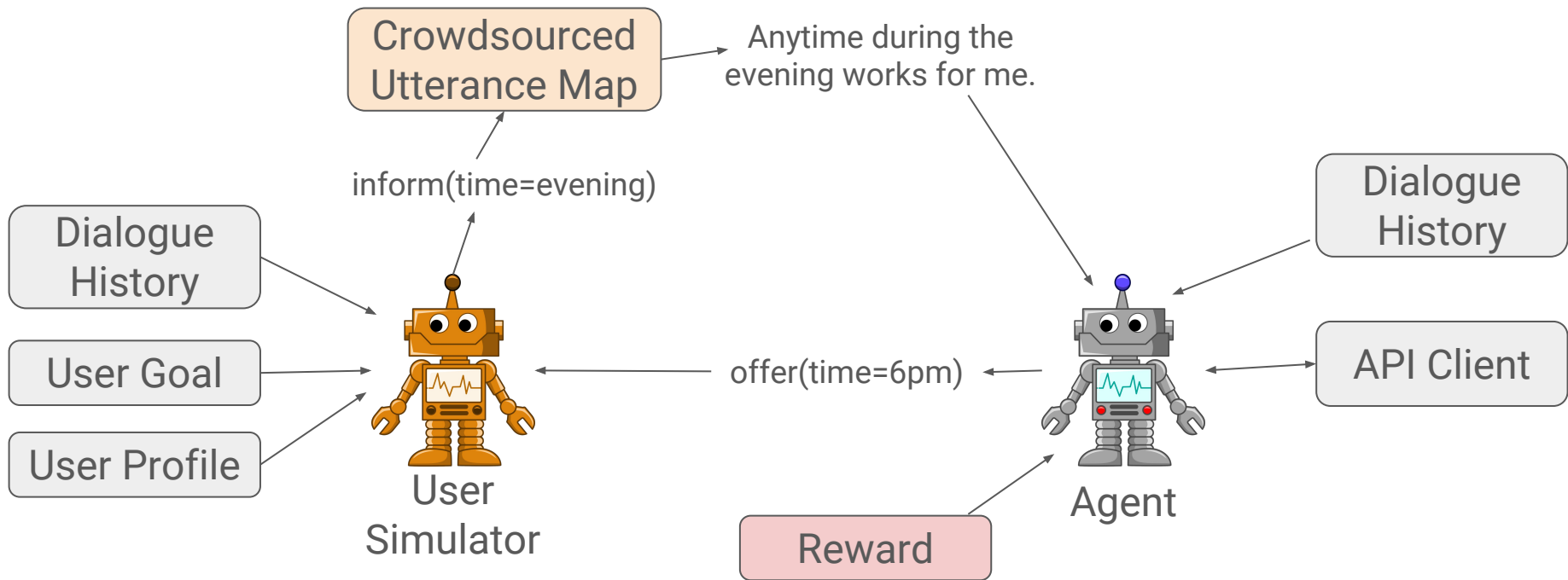
System actions

Policy Gradient Loss:

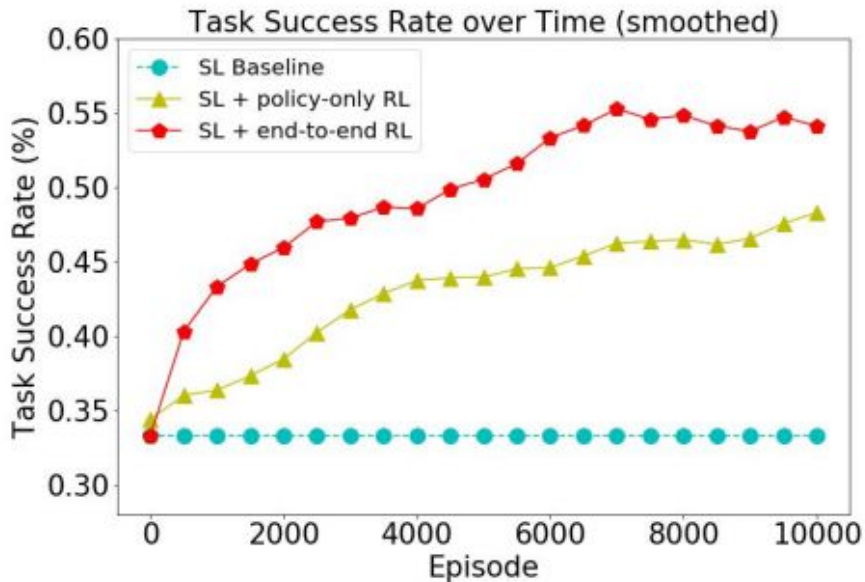
$$\nabla_{\theta} \mathbb{E}_{\theta} [R_k] = \mathbb{E}_{\theta_a} [\nabla_{\theta} \log \pi_{\theta}(a_k | s_k) R_k]$$

Bing Liu, Gokhan Tur, Dilek Hakkani-Tur, Pararth Shah, and Larry Heck. "End-to-End Optimization of Task-Oriented Dialogue Model with Deep Reinforcement Learning." Conversational AI Workshop, NIPS 2017.

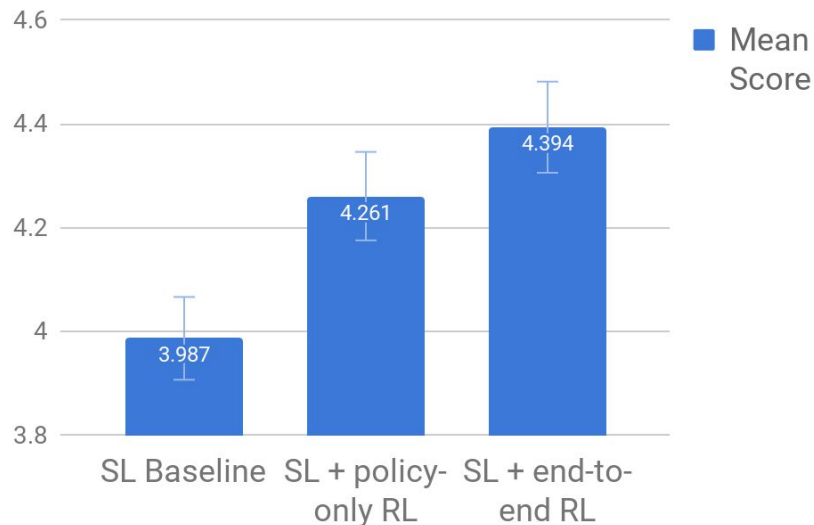
RL with User Simulator



Model evaluation



Human evaluation Turn scores from 1 to 5 Averaged over 100 dialogues



Bing Liu, Gokhan Tur, Dilek Hakkani-Tur, Pararth Shah, and Larry Heck. "End-to-End Optimization of Task-Oriented Dialogue Model with Deep Reinforcement Learning." Conversational AI Workshop, NIPS 2017.

Applications: Contextual language understanding

U: Find restaurants in Mountain View.

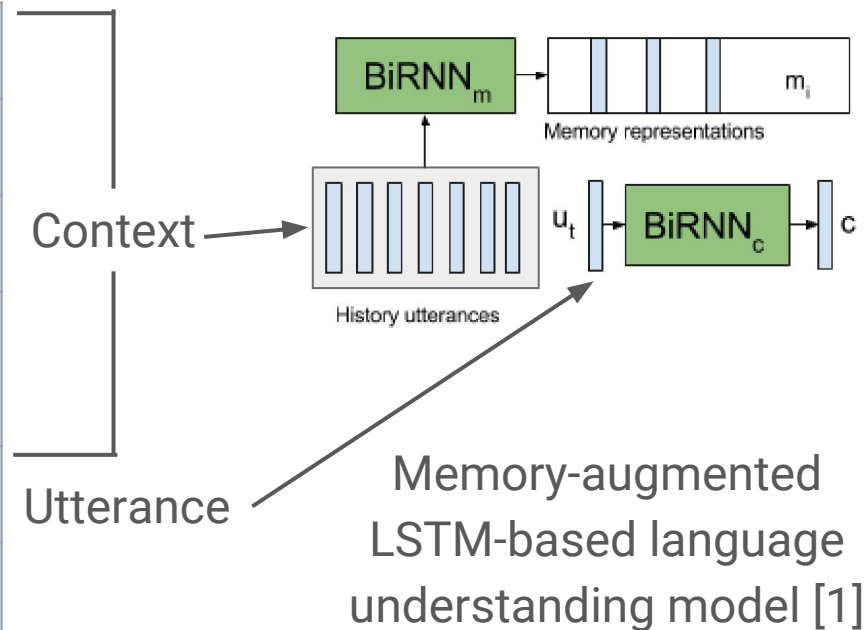
A: Cascad and Evvia are some options.

U: Does **the first one** have outdoor seating?

A: Yes, **Cascad** has outdoor seating. Should I book a table **there**?

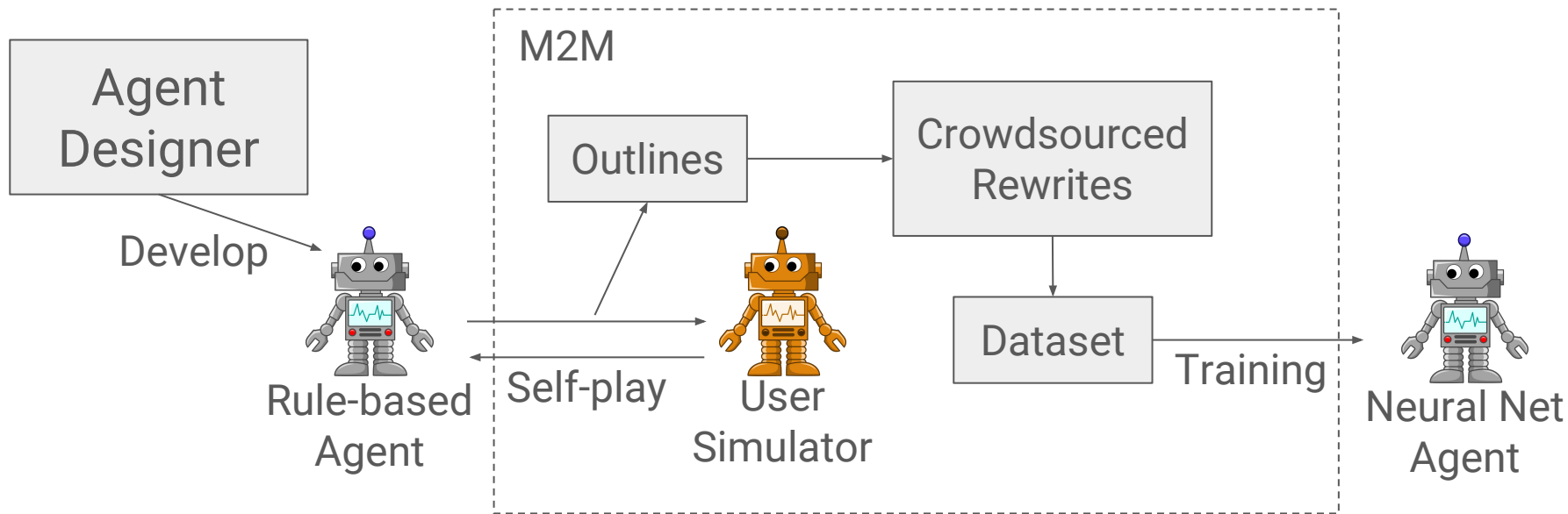
U: Yes, I am having **dinner** with a **friend**.

A: OK, there is a table for **two** at **7pm**.



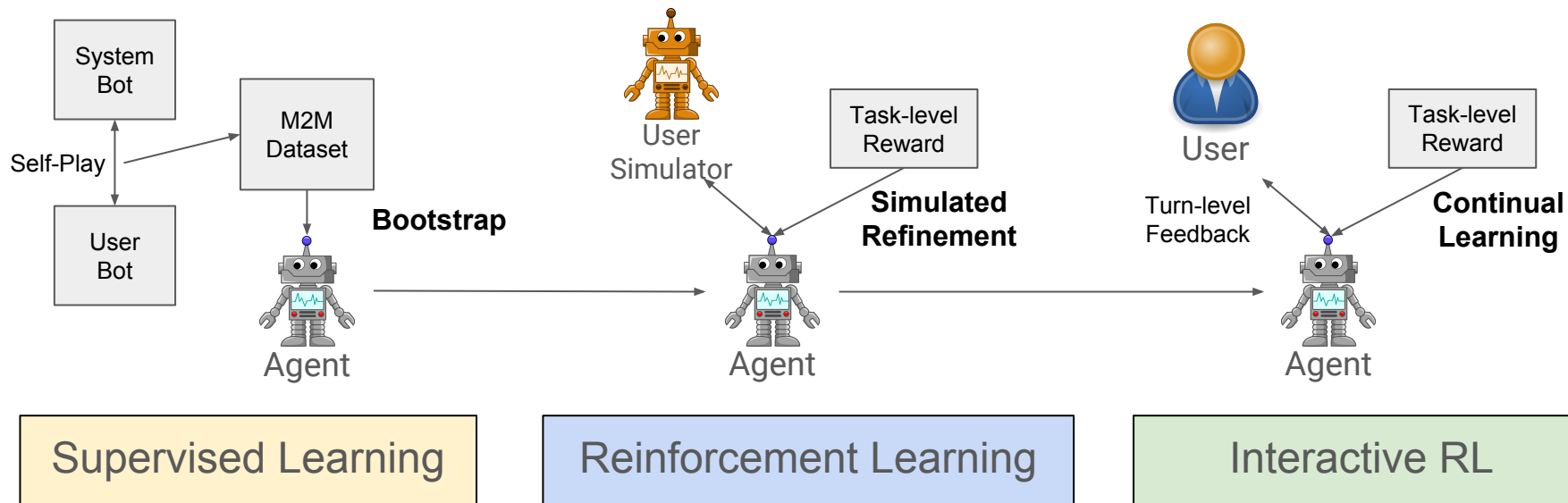
[1] Ankur Bapna, Gokhan Tur, Dilek Hakkani-Tur, and Larry Heck. "Sequential Dialogue Context Modeling for Spoken Language Understanding." In Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, pp. 103-114. 2017.

Distill expert dialogue policy into end-to-end NNet



Lifelong learning with human-in-the-loop RL

Multi-stage training of conversational agents



Pararth Shah, Dilek Hakkani-Tür, and Larry Heck. "Interactive reinforcement learning for task-oriented dialogue management." Deep Learning for Action and Interaction Workshop, NIPS 2016.

Thank you!

Pararth Shah

Google AI

: @PararthS