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
- Coreference
- Implicature
- Entailment
- 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

Intents:
- Weather queries
- Booking movie tickets
- Reserving restaurant tables
- ...

Agent Designer

Language Understanding + Dialogue Policy

Develop

Agent

Deploy

User

+ 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?

- Guaranteed coverage of critical user journeys
- Improve language understanding from data and experience
- Efficiently add new capabilities
- Flexibility
- Language complexity
- Scale
- Task complexity
Idea 1: Dialogue is a collaborative game

User Goal

User Behavior

User:

Dialogue Moves

Buy me a movie ticket

Sure, which movie?

Agent:

Web:

Reward
Idea 2: Separate task/discourse from language

**User**

**Agent**

### “Dialogue Outline”

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

...

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

Dialogue History
User Goal
User Profile

inform(time=evening)
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 real 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.
- Your paraphrased conversation needs to make sense on its own.

<table>
<thead>
<tr>
<th>Task</th>
<th>n_u</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistant: greetings</td>
<td>Hello</td>
</tr>
<tr>
<td>User: greetings AND buy movie tickets AND data is Wednesday</td>
<td>I would like to buy movie tickets for Wednesday.</td>
</tr>
<tr>
<td>Assistant: provide preference for: theatre name and time and number of tickets and movie</td>
<td>What theatre, movie, time, and number of tickets would you like?</td>
</tr>
<tr>
<td>User: unable to understand, must re-phrase your message</td>
<td>Sorry, I didn't understand. Could you repeat that?</td>
</tr>
<tr>
<td>Outline</td>
<td>Semantic annotations</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td><strong>Semantic annotations</strong></td>
<td>A: greeting()</td>
</tr>
<tr>
<td></td>
<td>U: inform(intent=book_movie, name=Inside Out, date=tomorrow, num_tickets=2)</td>
</tr>
<tr>
<td></td>
<td>A: ack() request(time)</td>
</tr>
<tr>
<td></td>
<td>U: inform(time=evening)</td>
</tr>
<tr>
<td></td>
<td>A: offer(theatre=Cinemark 16, time=6pm)</td>
</tr>
<tr>
<td></td>
<td>U: affirm() inform(intent=find_rest, meal=dinner, location=near theatre)</td>
</tr>
</tbody>
</table>
Combining automation & human intelligence

Machines Talking to Machines (M2M)

Task Schema

API Client

Scenarios

Self-Play (~5 mins)

Outlines

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:

(a) Task Specification

API Client:

Movies DB

Restaurants DB

(b) Scenario

book_movie:
- name=Inside Out
- theatre=don't care
- time=evening

reserve_restaurant:
- location=near theatre
- time=after the movie
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

Crowdsourced Rewrites (~2 hrs)

Natural Dialogues

Validation (~4 hrs)

Labelled Dataset

Training (~1 hr)

Scale to new tasks overnight (8hrs)
- With task schema and API client
- Task complexity
Dataset release

[GitHub Link]

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Schema</th>
<th>Train</th>
<th>Dev</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>M2M Restaurants</td>
<td>price_range, location, restaurant_name, category, num_people, date, time</td>
<td>1116</td>
<td>349</td>
<td>775</td>
</tr>
<tr>
<td>M2M Movies</td>
<td>theatre_name, movie, date, time, num_people</td>
<td>384</td>
<td>120</td>
<td>264</td>
</tr>
</tbody>
</table>

### Data quality evaluation

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

**Human evaluation**

![Human evaluation bar chart](chart.png)

Agent architecture

Policy Gradient Loss:
\[
\nabla_\theta \mathbb{E}_\theta [R_{k}] = \mathbb{E}_{\theta_a} [\nabla_\theta \log \pi_\theta (a_k | s_k) R_k]
\]

Goal slot values
\[
\min_\theta \sum_{k=1}^{K} - \left[ \sum_{m=1}^{M} \lambda_{lm} \log P(l_k^{m*} | U_{\leq k}, A_{< k}, E_{< k}; \theta) \right]
\]
\[+ \lambda_a \log P(a_k^* | U_{\leq k}, A_{< k}, E_{< k}; \theta) \]

System actions

RL with User Simulator

User Simulator

Crowdsourced Utterance Map

Anytime during the evening works for me.

Dialog History

User Goal

User Profile

API Client

Agent

Offer (time=6pm)

Reward

Inform (time=evening)
Model evaluation

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

**Applications:** Contextual language understanding

<table>
<thead>
<tr>
<th>Utterance</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>U: Find restaurants in Mountain View.</td>
<td>A: Cascal and Evvia are some options.</td>
</tr>
<tr>
<td>U: Does <strong>the first one</strong> have outdoor seating?</td>
<td>A: Yes, <strong>Cascal</strong> has outdoor seating. Should I book a table <strong>there</strong>?</td>
</tr>
<tr>
<td>U: Yes, I am having <strong>dinner</strong> with a <strong>friend</strong>.</td>
<td>A: OK, there is a table for <strong>two</strong> at <strong>7pm</strong>.</td>
</tr>
</tbody>
</table>

Distill expert dialogue policy into end-to-end NNet

- Agent Designer
  - Develop
  - Rule-based Agent

- M2M
  - Outlines
  - Self-play

- User Simulator

- Crowdsourced Rewrites

- Dataset

- Training
  - Neural Net Agent
Lifelong learning with human-in-the-loop RL

Multi-stage training of conversational agents

Thank you!