Personalization in Goal-oriented Dialog

Chaitanya K. Joshi, Fei Mi, Boi Faltings
Laboratory of Artificial Intelligence, EPFL, Switzerland

Poster at the Conversational AI Workshop at NIPS 2017
About me

CS Undergraduate, 3rd year @ NTU, Singapore

Exchange student @ EPFL, Switzerland

Intern @ SAP Machine Learning, Singapore
Conversational AI
Quick Background
Open-ended (chat-chat)

Slot-filling & Rule-based

Data/Corpus Driven

End-to-end Trained

Goal-oriented
Motivation

Personalization of a dialog system’s response based on who it is interacting with

- Real world conversations are influenced by characteristics and attributes of the speakers
- No open datasets for analyzing this problem!
Contributions

1. **New dataset** of goal-oriented dialogs influenced by speaker profiles
2. Modified **Memory Network architecture** to tackle personalization
3. Analysis of personalization as a **multi-task learning** problem
bAbI Dialog Tasks

Framework to test end-to-end trained dialogue systems in goal oriented conversation (Restaurant reservation)

- Restaurant KB - location, cuisine, price range, rating, etc; queried by API calls
- User can say something in 4 different ways, bot has a unique way to reply
- Evaluation based on per-response accuracy (Discriminative models)
Full dialog is broken down into subtasks which test crucial skills-

1. **Issuing API calls - dialog state tracking**
2. **Updating API calls - dialog state tracking**
3. **Displaying options - using KB facts**
4. **Providing information - using KB facts**
5. **Conducting a full dialog**
Personalized bAbI Dialog Tasks
by altering bot utterance patterns and KB entities
Given a user’s profile

in addition to the original goals of the bAbI tasks, the bot must:

1. Alter speech style based on gender and age
   - Forming associations between language concepts (formality, precision) and user attributes

2. Personalize suggestions based on dietary preferences, favorite food item, etc.
   - Perform reasoning, ranking and retrieval based on combination of user attributes and variable parameters from conversation history
User Profile
Gender: Male, Age: Elderly, Favorite Food Item: Shepherd's Pie

Hi!
Greetings Sir, what may I assist you with today?
I'd like to book a table for six people in an expensive price range with British food.
I will start the request now.
May I know where the restaurant should be located?
In London.
Excellent Sir, please give me a moment to provide you with options
api_call(British, London, Six, Expensive)
Actually I would prefer for four.
No.
I will modify your request, is there anything else to change?
No.
Excellent Sir, please give me a moment to provide you with options
api_call(British, London, Four, Expensive)
May I suggest this option: The Place
Let's do it!
Excellent, I will finalize your request.
Can you provide me the address?
You rock!
May I help you in any other way, Sir
No thanks.
It was a pleasure to be of help to you, Sir

User Profile
Gender: Male, Age: Young, Favorite Food Item: Fish and Chips

Hi!
Hey dude, what's up?!
I'd like to book a table for six people in an expensive price range with British food.
I'm on it!
Where should it be?
In London.
Ok, looking for options
api_call(British, London, Six, Expensive)
Actually I would prefer for four.
Cool! Anything else you want to update?
Ok, looking for options
api_call(British, London, Four, Expensive)

Is this one cool: The Place
Let's do it!
Cool, its done!
Can you provide me the address?
You rock!
Want anything else?
No thanks.
No problem!
Memory Networks for Dialog

At each turn of dialog, iteratively read from and reason over conversation history to produce a response
Generic Architecture

- Input Module
- Memory Module
- Output Module
- Attention
- Controller

Embeddings flow:
- Input Module ➔ Controller
- Memory Module ➔ Output Module
- Attention ➔ Output Module

Embeddings paths:
- Input Module ➔ Memory Module
- Memory Module ➔ Output Module
- Attention ➔ Output Module
Memory Module

\[
\{0.1, 0.7, 0.2\}
\]

Dot product + softmax

\[0.1\vec{m}_1 + 0.7\vec{m}_2 + 0.2\vec{m}_3\]

\[\vec{u}_2\]

Controller

\[\vec{u}_1\]

Answer

kitchen

1: Sam moved to garden
2: Sam went to kitchen
3: Sam drops apple there

Where is Sam?

Input story
Memory Network
with
Split Memory Architecture
## Experimental Results

Per-response accuracy:

<table>
<thead>
<tr>
<th>Task</th>
<th>Rule-based System</th>
<th>Supervised Embeddings</th>
<th>Memory Networks Standard</th>
<th>Memory Networks Split memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT1: Issuing API calls</td>
<td>100</td>
<td>84.37</td>
<td>99.83 (98.87)</td>
<td>85.66 (82.44)</td>
</tr>
<tr>
<td>PT2: Updating API calls</td>
<td>100</td>
<td>12.07</td>
<td>99.99 (99.93)</td>
<td>93.42 (91.27)</td>
</tr>
<tr>
<td>PT3: Displaying options</td>
<td>100</td>
<td>9.21</td>
<td>58.94 (58.71)</td>
<td>68.60 (68.56)</td>
</tr>
<tr>
<td>PT4: Providing information</td>
<td>100</td>
<td>4.76</td>
<td>57.17 (57.17)</td>
<td>57.17 (57.11)</td>
</tr>
<tr>
<td>PT5: Full dialog</td>
<td>100</td>
<td>51.6</td>
<td>85.10 (77.74)</td>
<td>87.28 (78.1)</td>
</tr>
</tbody>
</table>
Visualizing Attention

Predictions by standard Memory Networks

<table>
<thead>
<tr>
<th>Time</th>
<th>Locutor</th>
<th>Dialog History</th>
<th>Hop #1</th>
<th>Hop #2</th>
<th>Hop #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>male elderly</td>
<td>0.15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>User</td>
<td>hello</td>
<td>0.306</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Bot</td>
<td>greetings sir what may i assist you with today</td>
<td>0.008</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>User</td>
<td>may i have a table in a moderate price range with italian food for eight</td>
<td>0.536</td>
<td>0.999</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Bot</td>
<td>excellent sir i will start the request now</td>
<td>0</td>
<td>0</td>
<td>0.999</td>
</tr>
</tbody>
</table>

User input: <SILENCE>

Correct answer: may i know where the restaurant should be located

Predicted answer: may i know where the restaurant should be located

- Able to use profile information to personalize speech style
<table>
<thead>
<tr>
<th>Time</th>
<th>Locutor</th>
<th>Dialog History</th>
<th>Hop #1</th>
<th>Hop #2</th>
<th>Hop #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>User</td>
<td>female elderly</td>
<td>0.006</td>
<td>0.012</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Bot</td>
<td>hi</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>User</td>
<td>good day madam how could i assist you today</td>
<td>0.006</td>
<td>0.248</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Bot</td>
<td>may i have a table</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>User</td>
<td>thank you madam i shall start the reservation now</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Bot</td>
<td>&lt;SILENCE&gt;</td>
<td>0.064</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>User</td>
<td>could you tell me your preference on the type of cuisine</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>Bot</td>
<td>i love french food</td>
<td>0.307</td>
<td>0.103</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>User</td>
<td>could you tell me where the restaurant should be located</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>Bot</td>
<td>madrid please</td>
<td>0.265</td>
<td>0.282</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>User</td>
<td>would you mind telling me how many guests shall be at your table</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>Bot</td>
<td>for four please</td>
<td>0.225</td>
<td>0.076</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>User</td>
<td>would you mind telling me your price range</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>Bot</td>
<td>in a cheap price range please</td>
<td>0.11</td>
<td>0.29</td>
<td>0.999</td>
</tr>
<tr>
<td>15</td>
<td>User</td>
<td>thank you madam i shall provide you with options shortly</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

User input: <SILENCE>

Correct answer: api_call french madrid four cheap

Predicted answer: api_call french madrid four cheap

- Able to track dialog state
<table>
<thead>
<tr>
<th>Time</th>
<th>Locutor</th>
<th>Dialog History</th>
<th>Hop #1</th>
<th>Hop #2</th>
<th>Hop #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>female young non-veg pizza</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>User</td>
<td>resto_rome_moderate_italian_8stars_1 R_rating 8</td>
<td>0.0001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>User</td>
<td>resto_rome_moderate_italian_8stars_1 R_type veg</td>
<td>0.0001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>User</td>
<td>resto_rome_moderate_italian_8stars_1 R_speciality pizza</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>User</td>
<td>resto_rome_moderate_italian_3stars_1 R_location rome</td>
<td>0.879</td>
<td>0.999</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>User</td>
<td>resto_rome_moderate_italian_3stars_1 R_type veg</td>
<td>0.0001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>User</td>
<td>resto_rome_moderate_italian_2stars_2 R_rating 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>User</td>
<td>resto_rome_moderate_italian_2stars_2 R_type non-veg</td>
<td>0.0001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>46</td>
<td>User</td>
<td>resto_rome_moderate_italian_2stars_2 R_speciality pasta</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>74</td>
<td>User</td>
<td>hi</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>75</td>
<td>Bot</td>
<td>hey girl how is it going</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>76</td>
<td>User</td>
<td>may i have a table in rome with italian food in a moderate price range</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>77</td>
<td>Bot</td>
<td>be right back with your reservation</td>
<td>0</td>
<td>0</td>
<td>0.999</td>
</tr>
<tr>
<td>83</td>
<td>Bot</td>
<td>how about this one: resto_rome_moderate_italian_8stars_2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>84</td>
<td>User</td>
<td>no i don’t like that</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>85</td>
<td>Bot</td>
<td>sure finding something else</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>86</td>
<td>User</td>
<td>&lt;SILENCE&gt;</td>
<td>0.0003</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>87</td>
<td>Bot</td>
<td>how about this one: resto_rome_moderate_italian_3stars_2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>88</td>
<td>User</td>
<td>no i don’t like that</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>89</td>
<td>Bot</td>
<td>sure finding something else</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**User Input**

< SILENCE >

**Correct Answer**

how about this one: resto_rome_moderate_italian_8stars_1

**Predicted Answer**

what food are you looking for

- Unable to personalize reasoning/retrieval over KB facts
Split Memory Architecture

➢ Attention to user’s profile makes sense

➢ Can interpret facts and context from conversation history
What’s next

Dataset

Models

Evaluation
Thank you!

Questions?

Contact me

chaitjo.github.io
github.com/chaitjo
linkedin.com/chaitjo
chaitjo@gmail.com