Integrating Learning of Dialog Strategies and Semantic Parsing

Aishwarya Padmakumar, Jesse Thomason, Raymond Mooney

Department of Computer Science
The University of Texas at Austin
Natural Language Interaction with Robots
Understanding Commands

Go to Bob Brown’s office

Bring coffee for Alice
Goal of this work

- Dialog Management
- Semantic Parsing
Semantic Parsing

\[ S : \text{navigate}(\lambda x. (\text{office}(x) \land \text{owns}(\text{alice}, x))) \]

\[ \text{NP} : \lambda x. (\text{office}(x) \land \text{owns}(\text{alice}, x)) \]

\[ \text{NP}/\text{NP} : \lambda y. (\lambda x. (\text{office}(x) \land \text{owns}(y, x))) \]

\[ S/\text{NP} : \lambda x. \text{navigate}(x) \]

\[ \text{NP} : \text{alice} \]

\[ \text{go} \quad \text{to} \quad \text{Alice} \]

\[ N : \text{office} \]

\[ \text{‘s} \]

\[ \text{office} \]
Advantages of semantic parsing

Bring coffee for Bob Brown

Bob Brown’s office

Yes, you’re right

Complete command

Noun Phrase

Confirmation
Advantages of semantic parsing

Alice
Alice Ashcraft
Ms Alice Ashcraft

Alice’s office
Alice Ashcraft’s office
Ms Alice Ashcraft’s office
Clarification Dialogs

- Obtain specific missing information (Thomason et al., 2015)
- Weak supervision for improving parser (Artzi et al., 2011; Thomason et al., 2015)
Clarification Dialogs

- Obtain specific missing information (Thomason et al., 2015)
- Weak supervision for improving parser (Artzi et al., 2011; Thomason et al., 2015)
Dialog policy

TASK TO COMPLETE
Dave Daniel wants the item in slot 5.

| ROBOT | How can I help? |
| YOU   | please bring the item in slot 5 to dave daniel |
| ROBOT | What should I bring to dave daniel? |
| YOU   | a calander |
| ROBOT | I'm sorry, but I couldn't pinpoint what you meant by that. |
| ROBOT | What should I bring to dave daniel? |
| YOU   | a day planner |
| ROBOT | I'm sorry, but I couldn't pinpoint what you meant by that. |
| ROBOT | What should I bring to dave daniel? |
| YOU   | calendar |
| ROBOT | You want me to bring calendar to dave daniel? |
| YOU   | yes |
| ROBOT | I thought so |
| ROBOT | Happy to help |

Final Task
Dialog policy

- Mapping from dialog states to dialog actions
- Dialog state
  - Information from the dialog so far
  - Probability of each command being what the user intends
- Static policy
  - Difficult to update as parser changes
  - Optimal policy may not be obvious to hand-code
Background: Reinforcement Learning

- Reinforcement learning - learning for sequential tasks with only implicit feedback
- Partially Observable Markov Decision Process -

\[ \text{Agent} (b_{t-1}) \rightarrow \text{Agent} (b_t) \]

\[ \text{Environment} (s_t) \rightarrow \text{Environment} (s_{t+1}) \]

\[ a_t \]

\[ o_t \]
Clarification Dialogs as POMDPs

- States: What the user is trying to convey via the dialog (commands)
- Observations: Output of the natural language understanding component
- Actions: Dialog actions (confirming, requesting a parameter value)
Why is policy learning challenging?

Assumption:

Constant probability distribution

Our system:

Variable probability distribution
Why is policy learning challenging?

Assumption:
- Constant probability distribution
- Non-stationary Environment
Policy learning

- Kalman Temporal Differences (KTD) Q Learning (Geist and Pietquin, 2010) used to learn policy in non-stationary environment
- Dialog state tracking - HIS Model (Young et al., 2010)
- Policy learned over features extracted from belief state
Experiments - Mechanical Turk

**TASK TO COMPLETE**

Dave Daniel wants the item in slot 5.

<table>
<thead>
<tr>
<th>ROBOT</th>
<th>HOW CAN I HELP?</th>
</tr>
</thead>
<tbody>
<tr>
<td>YOU</td>
<td>PLEASE BRING THE ITEM IN SLOT 5 TO DAVE DANIEL</td>
</tr>
<tr>
<td>ROBOT</td>
<td>TO WHOM SHOULD I BRING SOMETHING?</td>
</tr>
<tr>
<td>YOU</td>
<td>DAVE DANIEL</td>
</tr>
<tr>
<td>ROBOT</td>
<td>WHAT ACTION DID YOU WANT ME TO TAKE INVOLVING DAVE DANIEL?</td>
</tr>
<tr>
<td>YOU</td>
<td>5</td>
</tr>
<tr>
<td>ROBOT</td>
<td>WHAT SHOULD I BRING TO DAVE DANIEL?</td>
</tr>
<tr>
<td>YOU</td>
<td>A CALENDAR</td>
</tr>
<tr>
<td>ROBOT</td>
<td>I'M SORRY, BUT I COULDN'T PINPOINT WHAT YOU MEANT BY THAT.</td>
</tr>
<tr>
<td>ROBOT</td>
<td>WHAT SHOULD I BRING TO DAVE DANIEL?</td>
</tr>
<tr>
<td>YOU</td>
<td>A DAY PLANNER</td>
</tr>
<tr>
<td>ROBOT</td>
<td>I'M SORRY, BUT I COULDN'T PINPOINT WHAT YOU MEANT BY THAT.</td>
</tr>
<tr>
<td>ROBOT</td>
<td>WHAT SHOULD I BRING TO DAVE DANIEL?</td>
</tr>
<tr>
<td>YOU</td>
<td>CALENDAR</td>
</tr>
<tr>
<td>ROBOT</td>
<td>YOU WANT ME TO BRING CALENDAR TO DAVE DANIEL?</td>
</tr>
<tr>
<td>YOU</td>
<td>YES</td>
</tr>
<tr>
<td>ROBOT</td>
<td>I THOUGHT SO</td>
</tr>
<tr>
<td>ROBOT</td>
<td>HAPPY TO HELP</td>
</tr>
</tbody>
</table>

**DIRECTORY**

**People:**

Alice Ashcraft; Secretary
Francis (“Frannie”) Foster
Robert (“Bob”) Brown
Carol Clark, PhD
Dave Daniel, PhD
George Green; Intern
Evelyn (“Eve”) Eckhart
Mallory Morgan; Director
Peggy Parker, PhD
Walter Ward; Supervisor

**Items available to robot:**

1. Coffee cup
2. Burger
3. Calendar
4. Day planner
5. Calendar
Hypotheses

• Combined parser and dialog learning is more useful than either alone.
• Changes in the parser need to be seen by the policy learner.
Systems

Parser and Dialog Learning - Full

Initial parser
Initial policy
Collect Dialogs
Update parser
Update policy
Final parser
Final policy

Parser and Dialog Learning - Batchwise

Initial parser
Initial policy
Collect Dialogs
Update parser
Update policy
Collect Dialogs
Collect Dialogs
Collect Dialogs
Update parser
Update policy
Final parser
Final policy
Final policy
Experiments - Metrics

● Successful dialog -
  ○ User does not exit dialog prematurely
  ○ Robot performs the correct action

● Objective metrics -
  ○ % successful dialogs
  ○ Dialog length - Average number of turns in successful dialogs

● Subjective metrics -
  ○ The robot understood me
  ○ The robot asked sensible questions
Results - Dialog Success

- Higher is better
- Parser learning is mostly responsible for improvement in dialog success rate
- Best system: parser and dialog learning - batchwise
Results - Dialog Length

- Lower is better
- Dialog learning is mostly responsible for lowering dialog length
- Best system: parser and dialog learning - batchwise
Results - user survey responses

- Higher is better
- Best system: parser and dialog learning - full
- Can be conflated with number of questions the system asked
Results - user survey responses

- Higher is better
- Best system: parser and dialog learning - batchwise
Conclusion

- Simultaneous training of a semantic parser and dialog policy using implicit feedback from dialogs.
  - Both batchwise > Parser learning, Dialog learning

- Changes in other components must be propagated to the policy via implicit feedback to reduce effects of non-stationary environment.
  - Both batchwise > Both full
Integrating Learning of Dialog Strategies and Semantic Parsing

Aishwarya Padmakumar, Jesse Thomason, Raymond Mooney

Department of Computer Science
The University of Texas at Austin