Opportunistic Active Learning for Grounding Natural Language Descriptions

Jesse Thomason, Aishwarya Padmakumar, Jivko Sinapov, Justin Hart, Peter Stone, Raymond J. Mooney

1 - Department of Computer Science, The University of Texas at Austin
2 - Department of Computer Science, Tufts University
Natural Language Interaction with Robots
Commands involving objects

Move the green bottle
Object Descriptions

“light empty yellow spiky container”
“a yellow pineapple”
“a shiny hollow empty container”
“wide aluminum silver can”

“it’s tall and squishy and red”
“pink foamy tall and cylindrical”
“a full neon green water bottle”
“a green water bottle that's heavy”
Learning during operation

Improving models during operation -

• Reduces the need for initial annotated data
• Aids domain adaptation of pretrained models
• Provides a way to handle novel objects and properties
Opportunistic Active Learning

- Active learning incorporated into a larger sequence of interactive tasks.
- Asking queries that may not be relevant to the current interaction but expected to be useful for future interactions.
Goal of this work

- Introduce opportunistic active learning where a system engaged in an interactive task uses active learning to acquire labels useful for future interactions.
- Demonstrate usefulness in grounding object descriptions with a real robot.
Object Descriptions - Non Visual Information

“light empty yellow spiky container”

“a yellow pineapple”

“a shiny hollow empty container”

“wide aluminum silver can”

“it's tall and squishy and red”

“pink foamy tall and cylindrical”

“a full neon green water bottle”

“a green water bottle that's heavy”
Object Descriptions

Difficult to annotate - Sparse Labels

“light empty yellow spiky container”
“a yellow pineapple”

“it's tall and squishy and red”
“pink foamy tall and cylindrical”
Active learning
Active Learning for Object Descriptions

The robot needs

- Object nearby
- To be interacting with a person

Heavy?
Active Learning for Object Descriptions
Active Learning for Object Descriptions
Active Learning for Object Descriptions
Prior Work: I Spy
Contrast with prior work

• No dedicated training phase.
• Incorporate active learning queries when performing interactive tasks.
• Robot is evaluated on end-task performance, not classifier accuracy.
Object Retrieval Task

Human: This is a yellow bottle with water filled in it.
Object Retrieval Task

- User describes an object in the active test set
- Robot needs to identify which object is being described
Object Retrieval Task

- User describes an object in the active test set
- Robot needs to identify which object is being described
Object Retrieval Task

- Robot can ask questions about objects on the sides to learn object attributes
Object Retrieval Task

- Robot can ask questions about objects on the sides to learn object attributes
Understanding descriptions

• The robot tokenizes the description and treats all words except stopwords as a predicate to learn a binary classifier for.

• The robot these classifiers from interactions and uses them to identify the object being described.
Multimodal Classifiers

The robot used vision, audio and haptic information from manipulating the objects to learn classifiers for object properties.
Querying a label for an object

Robot: Would you use the word "bottle" to describe this object?
Querying a label for an object

• A yes/no question about whether a predicate applies to an object
• Predicate and object chosen using multi-label active learning, favouring predicates with classifiers of low confidence.
• Classifier prediction confidence is measured using Cohen’s Kappa estimated by leave-one-out cross validation
Querying a label for an object

$$o_{\text{min}}(p) = \arg\min_{o \in O_{tr}} (\kappa(p, o))$$

Object with least confidence for $p$

Confidence in prediction of $p$ for object $o$

$$\text{prob}(p) = \frac{1 - \kappa(p, o_{\text{min}}(p))}{\sum_{q \in P \setminus \{p\}} 1 - \kappa(q, o_{\text{min}}(q))}.$$ 

Probability of choosing $p$

Confidence in $p$ for the object with lowest confidence
Querying for positive example

Robot: Can you show me an object that you would describe as “yellow”?
Querying for positive example

• Request for positive example for a predicate
• Needed because most labels are sparse
• Predicate sampled uniformly among those without sufficient data to fit a classifier
Sample Interaction

<table>
<thead>
<tr>
<th>USER</th>
<th>This is a yellow bottle with water filled in it</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROBOT</td>
<td>Can you show me an object you would describe as “yellow”?</td>
</tr>
<tr>
<td>USER</td>
<td>(picking up a yellow object) This one</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td>ROBOT</td>
<td>Thanks! (points to a test object) Is this the object you were referring to?</td>
</tr>
<tr>
<td>USER</td>
<td>Yes</td>
</tr>
</tbody>
</table>
### On-topic and Off-topic questions

<table>
<thead>
<tr>
<th>On-Topic</th>
<th>Off-Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>USER</strong></td>
<td><strong>USER</strong></td>
</tr>
<tr>
<td>This is a <strong>yellow</strong> bottle with water filled in it</td>
<td>This is a <strong>yellow</strong> bottle with water filled in it</td>
</tr>
<tr>
<td><strong>ROBOT</strong></td>
<td><strong>ROBOT</strong></td>
</tr>
<tr>
<td>Can you show me an object you would describe as “<strong>yellow</strong>”?</td>
<td>Can you show me an object you would describe as “<strong>red</strong>”?</td>
</tr>
<tr>
<td><strong>USER</strong></td>
<td><strong>USER</strong></td>
</tr>
<tr>
<td>(picking up a yellow object) This one</td>
<td>(picking up a red object) This one</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td><strong>ROBOT</strong></td>
<td><strong>ROBOT</strong></td>
</tr>
<tr>
<td>Thanks! (points to a test object) Is this the object you were referring to?</td>
<td>Thanks! (points to a test object) Is this the object you were referring to?</td>
</tr>
<tr>
<td><strong>USER</strong></td>
<td><strong>USER</strong></td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Opportunistic Active Learning

• Asking locally convenient questions during an interactive task
• Questions may not be useful for the current interaction (off-topic) but expected to help future interactions
Opportunistic Active Learning - Challenges

Some other object might be a better candidate for the question

Purple?
Opportunistic Active Learning - Challenges

The question interrupts another task and may be seen as unnatural

Find the yellow bottle with water filled in it

Can you show me an object you would describe as “red”? 
Opportunistic Active Learning - Challenges

The information needs to be useful for a future task.

Red?
Experiments

Hypothesis: Opportunistic active learning helps a robot perform better at understanding natural language descriptions of objects.
Experimental Conditions

• Baseline vs Inquisitive condition
• Common steps
  – Get a description to identify
  – Ask questions about objects in active training set
  – Guess object from active test set
Experimental Conditions

Inquisitive condition -

• More talkative - Ask more questions before guessing than baseline
• Less task oriented - Can ask off-topic questions
Experiments

Hypothesis: Robot would perform better at retrieving objects in the inquisitive condition, without being perceived as being too frustrating.
Experiments

Objects used - 4 folds of 8 objects each
Results
Results

The robot seemed to understand my descriptions.

![Bar chart showing average Likert response per round for Baseline and Inquisitive conditions.]

- Baseline
- Inquisitive

Average Likert Response

<table>
<thead>
<tr>
<th>Round</th>
<th>Baseline</th>
<th>Inquisitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>
Results

The robot asked too many questions.

<table>
<thead>
<tr>
<th>Round</th>
<th>Average Likert Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0 (Baseline)</td>
</tr>
<tr>
<td>2</td>
<td>1.1 (Baseline)</td>
</tr>
<tr>
<td>3</td>
<td>1.0 (Baseline)</td>
</tr>
</tbody>
</table>
Results

It was fun to interact with the robot.

![Bar chart](chart.png)

- Baseline
- Inquisitive

Average Likert Response vs. Round

- Round 1
- Round 2
- Round 3
Results

I would use a robot like this to get objects for me in another room.

<table>
<thead>
<tr>
<th>Round</th>
<th>Baseline</th>
<th>Inquisitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
Conclusion

We introduce the setting of opportunistic active learning, where a system engaged in a task, queries for labels using active learning to improve performance in future tasks.
Conclusion

• Inquisitive robot performs better at understanding object descriptions.
• This setting makes the users find the robot more comprehending, fun and usable in a real-world setting.
Opportunistic Active Learning for Grounding Natural Language Descriptions

Jesse Thomason,\textsuperscript{1} Aishwarya Padmakumar,\textsuperscript{1} Jivko Sinapov,\textsuperscript{2} Justin Hart,\textsuperscript{1} Peter Stone,\textsuperscript{1} Raymond J. Mooney \textsuperscript{1}

\textsuperscript{1} - Department of Computer Science, The University of Texas at Austin
\textsuperscript{2} - Department of Computer Science, Tufts University
User Survey

Users rated the following statements on a Likert scale of 0-4 after interacting with the robot.
Asking Questions

Robot: Would you use the word "bottle" to describe this object?
Asking Questions

Robot: Can you show me an object that you would describe as “yellow”?
Experimental Conditions

- In baseline condition, the robot can only ask about “yellow”, “bottle”, “water”, “filled”
- In inquisitive condition, the robot can ask about any predicate it knows, possible “red” or “heavy”
## Experiments

<table>
<thead>
<tr>
<th>Round</th>
<th>Fold used for Active Training Set</th>
<th>Fold used for Active Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3 (test)</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Object Descriptions

Difficult to annotate - Objects may not be canonical examples of labels

“light empty yellow spiky container”

“a yellow pineapple”

“a full neon green water bottle”

“a green water bottle that's heavy”
Object Descriptions

Difficult to annotate - Descriptions may be relative to other objects

“light empty yellow spiky container”

“a yellow pineapple”

“a full neon green water bottle”

“a green water bottle that's heavy”