Learning a Policy for Opportunistic Active Learning

Aishwarya Padmakumar, Peter Stone, Raymond J. Mooney

Department of Computer Science
The University of Texas at Austin
Natural Language Interaction with Robots
Objects in Human Environments
Objects in Human Environments

- Diverse
Objects in Human Environments

- Diverse
- Transient
Objects in Human Environments

- Diverse
- Transient
- Described using diverse perceptual properties

“light empty yellow spiky container”
“a yellow pineapple”
“a full neon green water bottle”
“a green water bottle that’s heavy”
Understanding Object Descriptions

Robots need be able to

- Ground language in perception.
- Handle novel perceptual concepts during operation.
Opportunistic Active Learning
(Thomason et al., CoRL 2017)

- A framework for incorporating active learning queries into test time interactions.
- Demonstrated improvement in learning novel perceptual concepts to ground natural language descriptions of objects.
Goal of this Work

Learning a dialog policy for an interactive object retrieval task.
Task Setup

- Based on task from prior work (Thomason et al., CoRL 2017)
- Our contribution - Setting it up in simulation using the Visual Genome dataset.
# Task Setup

<table>
<thead>
<tr>
<th>Active Training Set</th>
<th>Dialog</th>
<th>Active Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Test_1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test_2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test_3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test_4</td>
</tr>
</tbody>
</table>
## Task Setup

<table>
<thead>
<tr>
<th>Active Training Set</th>
<th>Dialog</th>
<th>Active Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robot: Describe the object I should find.</td>
<td>Human: A white umbrella.</td>
<td>Test_1 Test_2 Test_3 Test_4</td>
</tr>
</tbody>
</table>

**Target Description**

**Target**
### Task Setup

<table>
<thead>
<tr>
<th>Active Training Set</th>
<th>Dialog</th>
<th>Active Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train_1</td>
<td>Robot: Describe the object I should find.</td>
<td>Test_1</td>
</tr>
<tr>
<td>Train_2</td>
<td>Human: A white umbrella.</td>
<td>Test_2</td>
</tr>
<tr>
<td>Train_3</td>
<td></td>
<td>Test_3</td>
</tr>
<tr>
<td>Train_4</td>
<td></td>
<td>Test_4</td>
</tr>
<tr>
<td>Train_5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Robot: Describe the object I should find.

Human: A white umbrella.
### Task Setup

<table>
<thead>
<tr>
<th>Active Training Set</th>
<th>Dialog</th>
<th>Active Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train_1</td>
<td>Robot: Describe the object I should find.</td>
<td>Test_1</td>
</tr>
<tr>
<td>Train_4</td>
<td>Human: A white umbrella.</td>
<td>Test_2</td>
</tr>
<tr>
<td>Train_2</td>
<td>Robot: Is there something in Train_6 that can be described as yellow?</td>
<td>Test_3</td>
</tr>
<tr>
<td>Train_3</td>
<td>Human: No.</td>
<td>Test_4</td>
</tr>
<tr>
<td>Train_6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Label Query**
<table>
<thead>
<tr>
<th>Active Training Set</th>
<th>Dialog</th>
<th>Active Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train_1</td>
<td>Robot: Describe the object I should find.</td>
<td>Test_1</td>
</tr>
<tr>
<td>Train_2</td>
<td>Human: A white umbrella.</td>
<td>Test_2</td>
</tr>
<tr>
<td>Train_3</td>
<td>Robot: Is there something in Train_6 that can be described as yellow?</td>
<td>Test_3</td>
</tr>
<tr>
<td>Train_4</td>
<td>Human: No.</td>
<td>Test_4</td>
</tr>
<tr>
<td>Train_5</td>
<td>Robot:</td>
<td></td>
</tr>
<tr>
<td>Train_6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Opportunistic Query**
### Active Training Set

<table>
<thead>
<tr>
<th>Train_1</th>
<th>Train_4</th>
<th>Train_2</th>
<th>Train_5</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
</tbody>
</table>

### Active Test Set

<table>
<thead>
<tr>
<th>Test_1</th>
<th>Test_2</th>
<th>Test_3</th>
<th>Test_4</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
</tbody>
</table>

### Dialog

**Robot:** Describe the object I should find.
**Human:** A white umbrella.
**Robot:** Is there something in Train_6 that can be described as yellow?
**Human:** No.

**Robot:** Can you show me an image with something that can be described as white?

**Human:** Train_1

---

Example Query
## Task Setup

<table>
<thead>
<tr>
<th>Active Training Set</th>
<th>Dialog</th>
<th>Active Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train_1</td>
<td><strong>Robot</strong>: Describe the object I should find.</td>
<td>Test_1</td>
</tr>
<tr>
<td></td>
<td><strong>Human</strong>: A white umbrella.</td>
<td>Test_2</td>
</tr>
<tr>
<td>Train_2</td>
<td><strong>Robot</strong>: Is there something in Train_6 that can be described as yellow?</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Human</strong>: No.</td>
<td>Test_3</td>
</tr>
<tr>
<td>Train_3</td>
<td><strong>Robot</strong>: Can you show me an image with something that can be described as white?</td>
<td>Test_4</td>
</tr>
<tr>
<td>Train_4</td>
<td><strong>Human</strong>: Train_1</td>
<td></td>
</tr>
<tr>
<td>Train_5</td>
<td><strong>Robot</strong>: My guess is Test_4</td>
<td></td>
</tr>
<tr>
<td>Train_6</td>
<td><strong>Human</strong>: Correct</td>
<td></td>
</tr>
<tr>
<td>Train_7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train_8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Robot: My guess is Test_4

Human: Correct
Goal of the Task

• Learn to maximize the fraction of successful guesses across interactions by
  – Learning when to ask queries, and when to stop and guess.
  – Learning to choose between different possible queries.
Active Learning
Why Opportunistic Queries?

- Robot may have good models for on-topic concepts.
- No useful on-topic queries.
- Some off-topic concepts may be more important because they are used in more 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 a green bottle.

Would you use the word “metallic” to describe this object?
Opportunistic Active Learning - Challenges

The information needs to be useful for a future task.

Red?
Grounding Model

A white umbrella $\rightarrow$ \{white, umbrella\}

↓

Pretrained CNN

SVM $\rightarrow$ white/ not white

SVM $\rightarrow$ umbrella/ not umbrella
Grounding Model

\[ o_{guess} = \arg\max_{o \in O_A} \sum_{i=1}^{k} d(p_i, o) * C(p_i) \]

Active Test Set
Grounding Model

\[ o_{guess} = \arg\max_{o \in O_A} \sum_{i=1}^{k} d(p_i, o) \times C(p_i) \]

- Active Test Set
- Classifier decision in \{-1, 1\}
Grounding Model

\[ o_{\text{guess}} = \operatorname{argmax}_{o \in O_A} \sum_{i=1}^{k} d(p_i, o) \times C(p_i) \]

- Active Test Set
- Classifier decision in \{-1, 1\}
- Classifier confidence in \(0, 1\)
Grounding Model

\[ o_{guess} = \arg\max_{o \in O_A} \sum_{i=1}^{k} d(p_i, o) \times C(p_i) \]

- **Active Test Set**
- Summed over predicates in description
- Classifier decision in \{-1, 1\}
- Classifier confidence in (0, 1)
Active Learning

• Agent starts with no classifiers.
• Labeled examples are acquired through questions and used to train the classifiers.
• Agent needs to learn a policy to balance active learning with task completion.
Modelling the Dialog as an MDP

Dialog Agent

Reward:
- +100 for correct guess
- -100 for incorrect guess
- -1 per to shorten dialogs

User

State:
- Target description
- Active Train and test objects
- Agent’s perceptual classifiers

Action:
- Label query
- Example Query
- Guess
Challenges

• What information about classifiers should be represented?
  – Features based on active learning metrics

• Variable number of queries and classifiers
  – Create features for state-action pairs

• Large action space
  – Sample a beam of promising queries
Feature Groups

• Query features - Active learning metrics used to determine whether a query is useful. Examples -
  – Current estimated F1 of classifier
  – Margin of object for classifier (for label query)
Feature Groups

• Guess features - Features that use the predictions and confidences of classifiers to determine whether a guess will be correct.

Examples -
  – Highest score among regions in the active test set.
  – Average estimated F1 of classifiers of concepts in description
Experiment Setup

- Policy learning using REINFORCE.
- Baseline - A hand-coded dialog policy that asks a fixed number of questions selected using the same sampling distribution.
Experiment Phases

- Initialization - Collect experience using the baseline to initialize the policy.
- Training - Improve the policy from on-policy experience.
- Testing - Policy weights are fixed, and we run a new set of interactions, starting with no classifiers, over an independent test set with different predicates.
Results

- Systems evaluated on dialog success rate and average dialog length.
- We prefer high success rate and low dialog length (top left corner)
Results
Learned policy is more successful than the baseline, while also using shorter dialogs on average.
If we ablate either group of features, the success rate drops considerably but dialogs are also much shorter.

In both cases, the system chooses to ask very few queries.
Summary

- We can learn a dialog policy that learns to acquire knowledge of predicates through opportunistic active learning.
- The learned policy is more successful at object retrieval than a static baseline, using fewer dialog turns on average.
Learning a Policy for Opportunistic Active Learning

Aishwarya Padmakumar, Peter Stone, Raymond J. Mooney

Department of Computer Science
The University of Texas at Austin
Policy Representation

\[ \pi(a | s; \theta) = \frac{e^{\theta^T f(s,a)}}{\sum_{a'} e^{\theta^T f(s,a')}} \]
Query Features

- Does the concept have a classifier?
- Current estimated F1 of the classifier
- Fraction of previous dialogs in which the predicate has been used, and the agent’s success rate in these.
- Is the query opportunistic?
Query Features

- Margin of object
- Density of object
- Fraction of k nearest neighbours of the object which are unlabelled
Guess Features

- Lowest, highest, second highest, and average estimated F1 among classifiers of concepts in the description.
- Highest score among regions in the active test set, and the differences between this and the second highest, and average scores respectively.
- An indicator of whether the two most confident classifiers agree on the decision of the top scoring region.
Sampling distribution