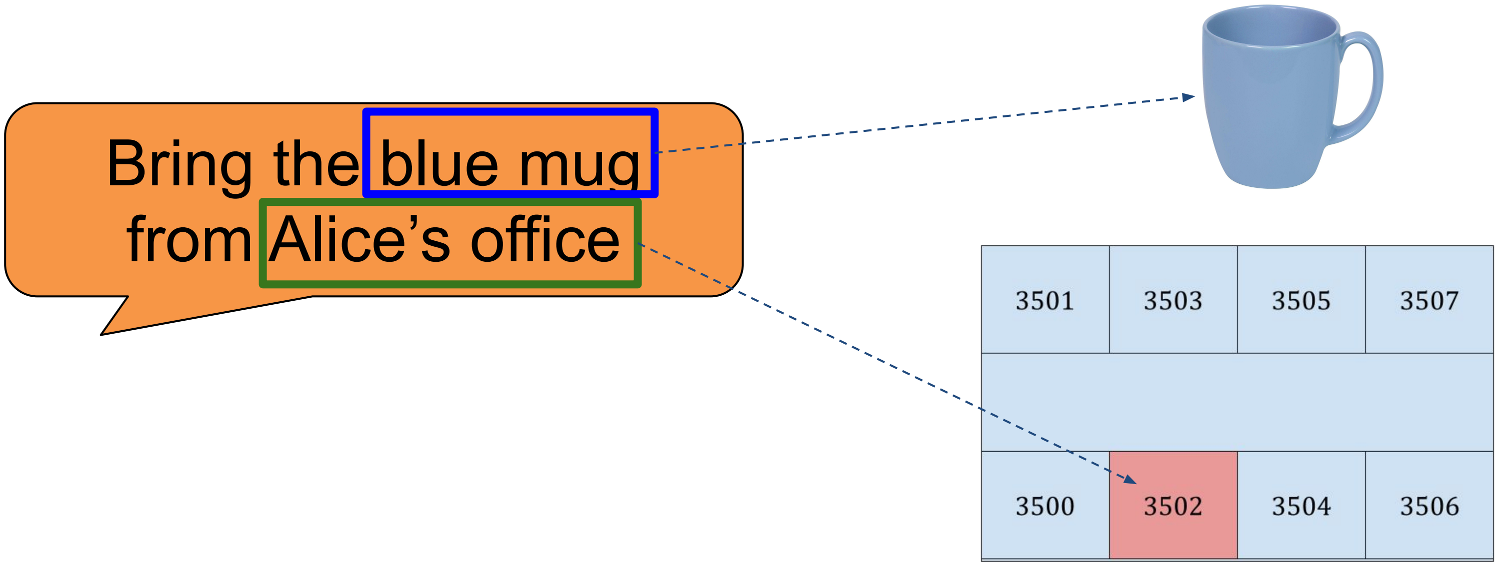
Dialog as a Vehicle for Lifelong Learning of Grounded Language Understanding Systems

Aishwarya Padmakumar

Doctoral Dissertation Defense

Grounded Language Understanding

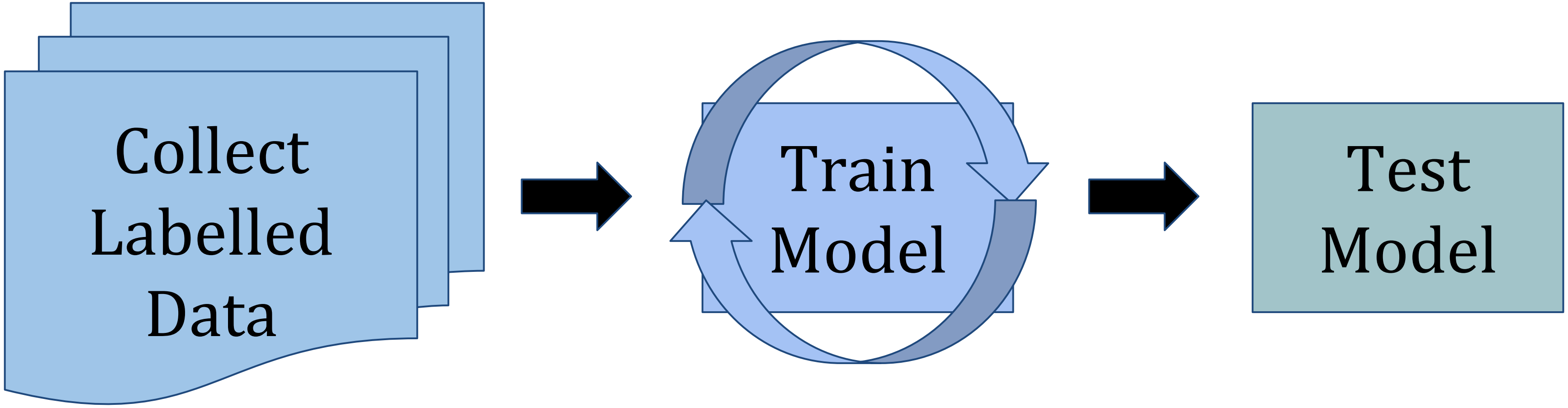
Mapping natural language to real-world entities



Applications in Service Robotics

Bring the blue mug
from Alice’s office

Standard Supervised Learning Pipeline



Sources of Imperfect Understanding

- Domain shift:
  - Train:  Test: 
- Missing domain specific knowledge
  - Alice’s office is missing in the directory
  - There is no category for mugs in the object detector.

Dialog - Clarification

Bring the blue mug from Alice's office

bring(,●)

Dialog - Clarification

Bring the blue mug from Alice's office

Where should I bring a blue mug from?

Alice Ashcraft's office

I should bring a blue mug from 3502?

Yes

Dialog - Improve Models

Bring the blue mug from Alice's office

Where should I bring a blue mug from?

Alice Ashcraft's office

I should bring a blue mug from 3502?

Yes

Alice's office

$\asymp$

Alice Ashcraft's office

$\asymp$

3502

Dialog - Acquiring Labels

Bring the blue mug from Alice's office

Blue?

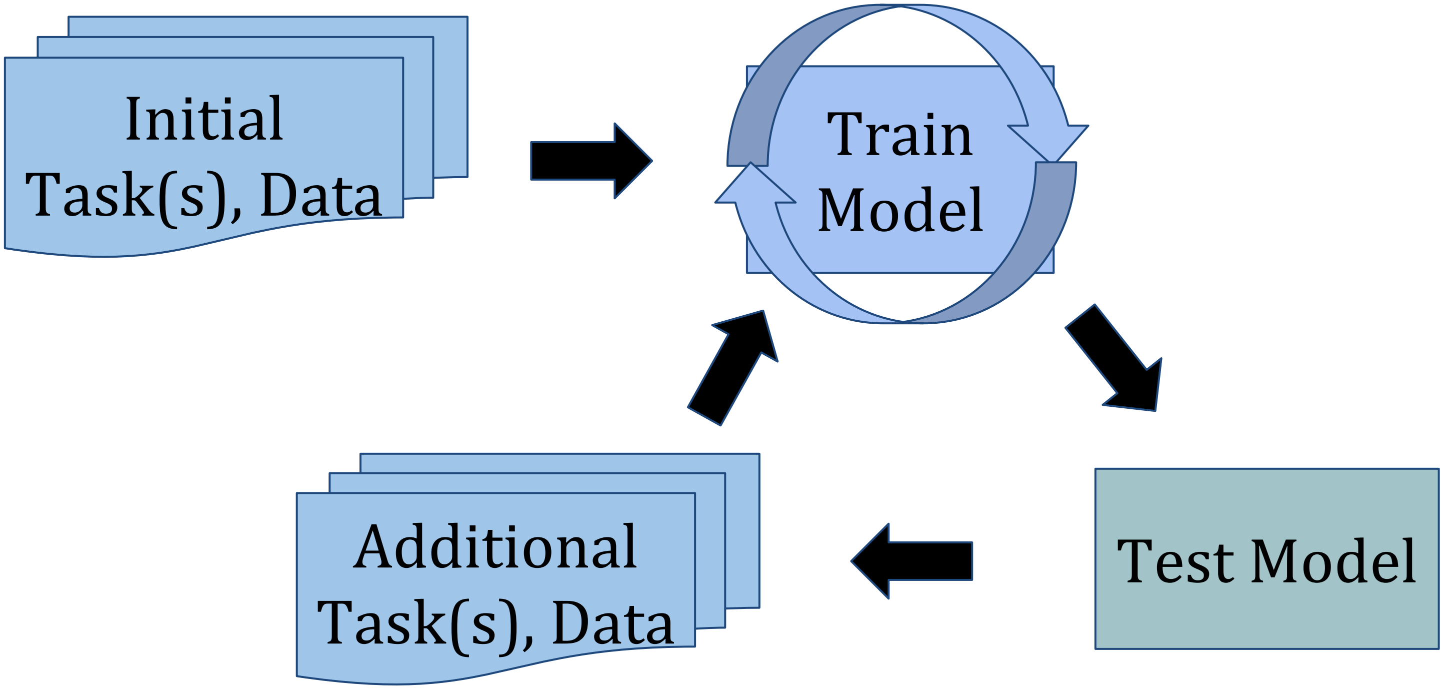
Dialog - Acquiring Labels

Bring the blue mug from Alice's office

Would you use the word "blue" to refer to this object?

Yes

Lifelong Learning



Lifelong Learning

Lifelong learning can make models more

- Generalizable - adapt to a variety of test data distributions
- Versatile - same model can be shared between multiple tasks, that are not necessarily pre-defined

Dialog as a Vehicle for Lifelong Learning

- Lifelong learning systems assume that additional labelled data can be obtained from test time usage.
- Dialog systems interact with users by design - interactions can be leveraged to obtain labelled data.

My Work

Designing dialog interactions to improve grounded language understanding systems and enabling them to perform lifelong learning.

Outline

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Outline

Pre-proposal Work

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Outline

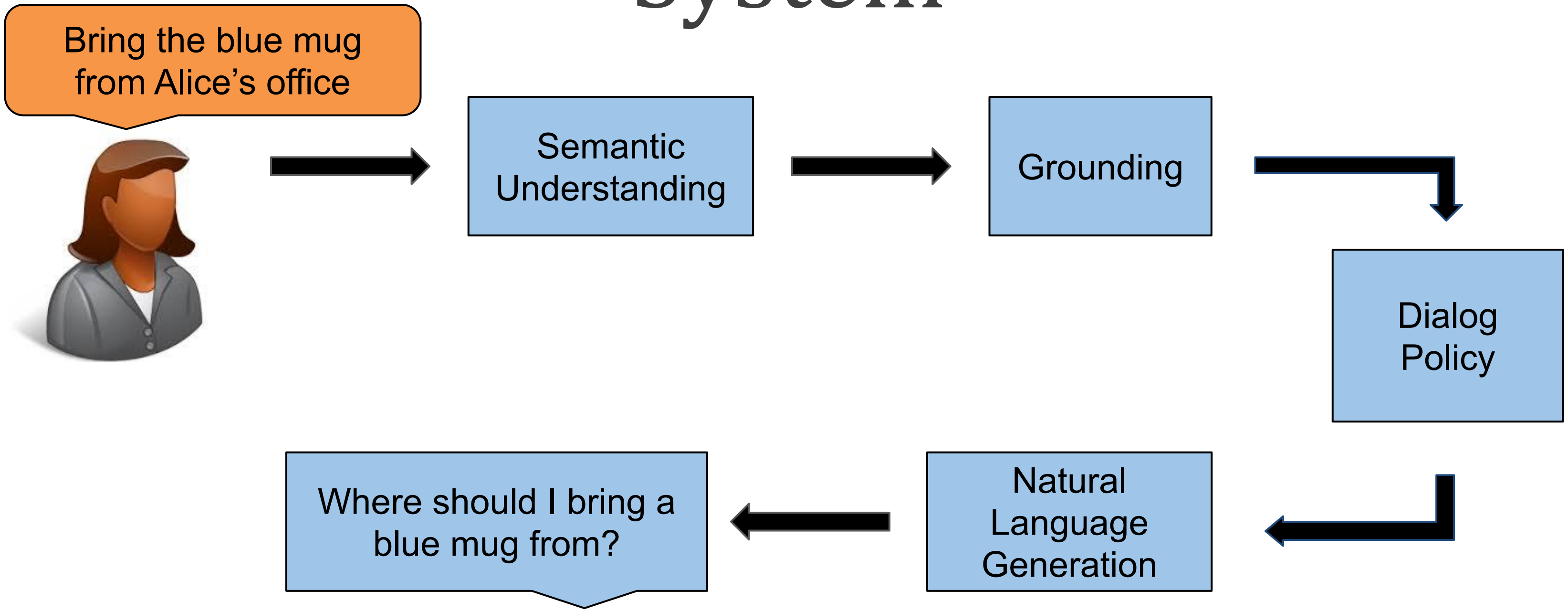
- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- **Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)**
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Post-proposal Work

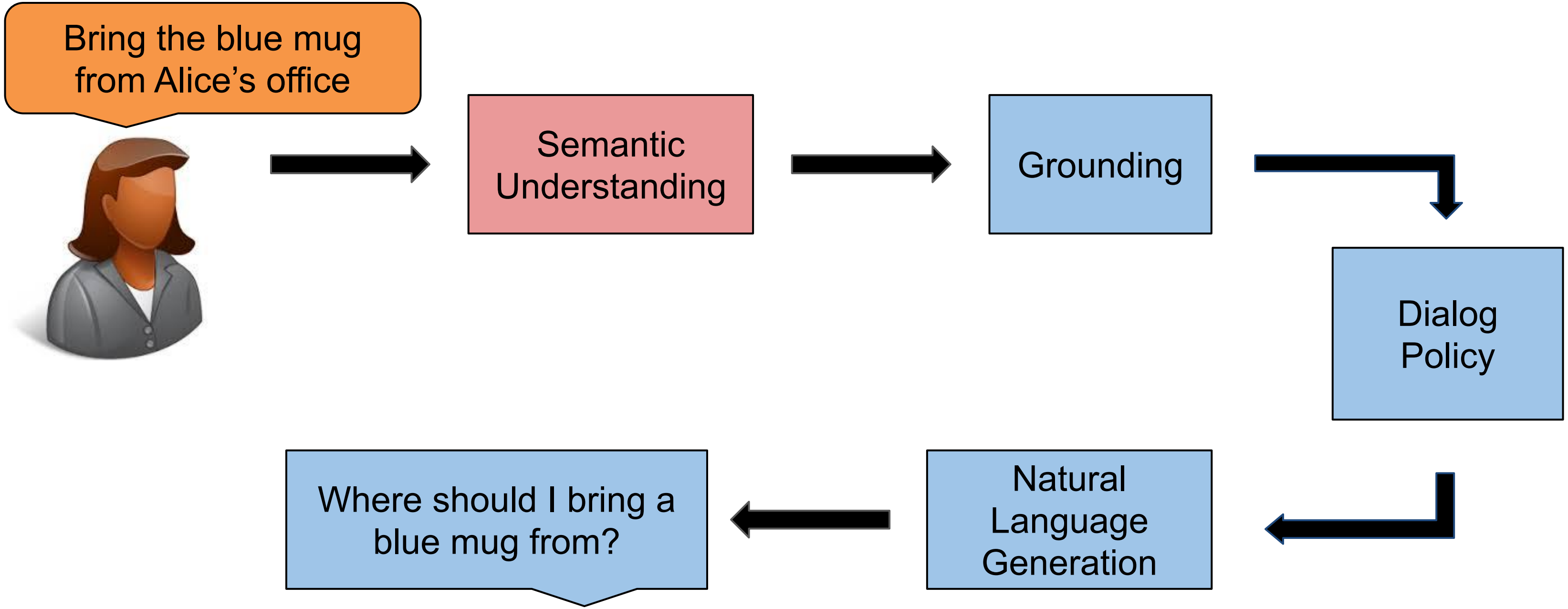
Outline

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Background: Parts of a Dialog System



Background: Semantic Understanding



Background: Semantic Understanding

Convert natural language into a machine understandable representation

Background: Semantic Understanding

Convert natural language into a machine understandable representation

Bring the blue mug from Alice's office

$$\mathtt{bring}(\lambda x.(\mathtt{blue}(x) \wedge \mathtt{mug}(x)), \\ \mathtt{the}(\lambda y.(\mathtt{office}(y) \wedge \mathtt{owns}(\mathtt{alice}, y))))$$

Semantic parsing -

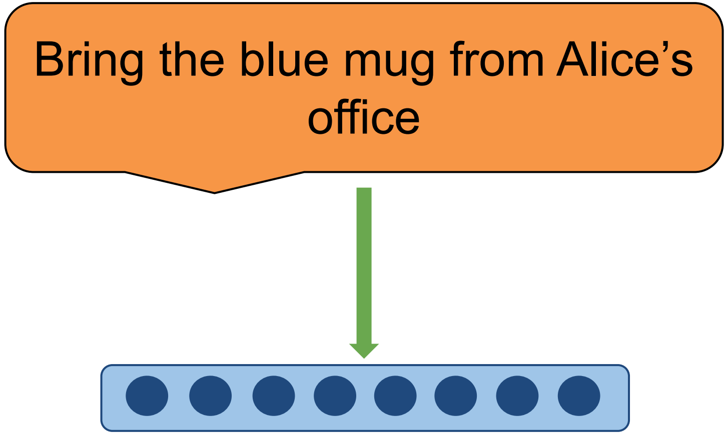
- Converts language to a structured meaning representation
- Compositionality - meaning of "*blue mug*" from meaning of "*blue*" and meaning of "*mug*"

Background: Semantic Understanding

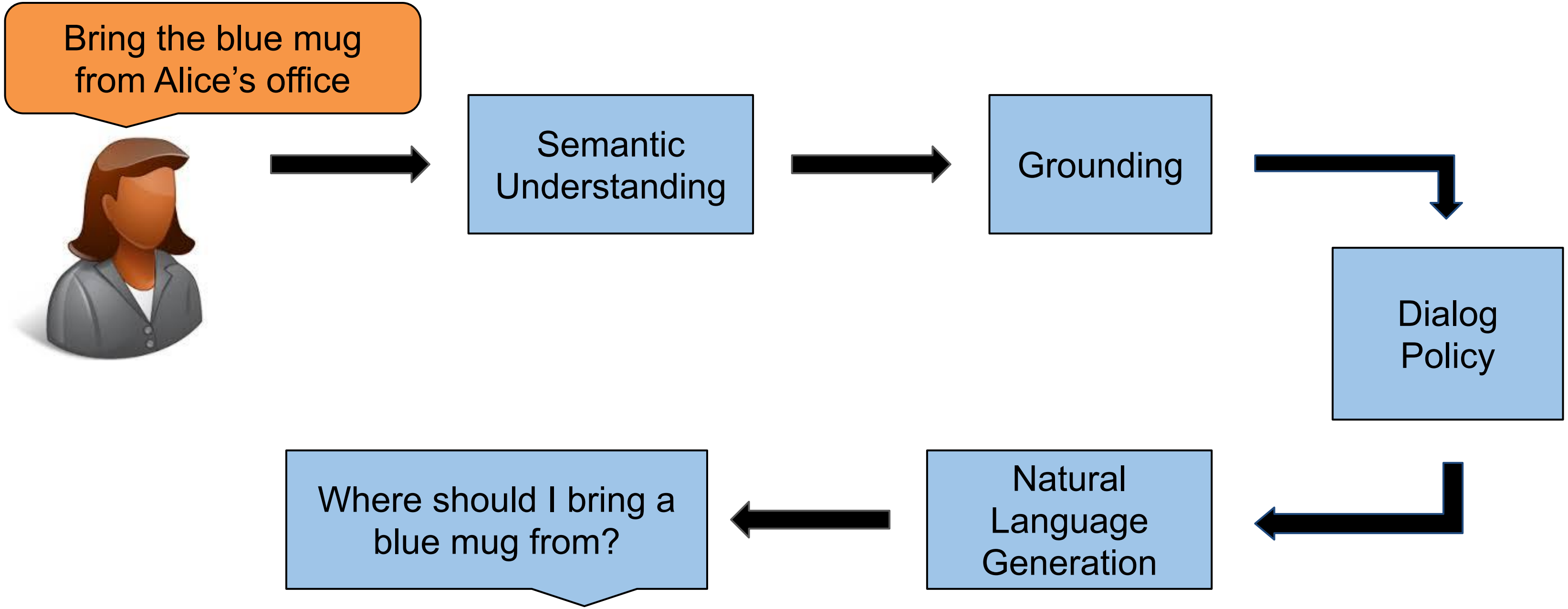
Convert natural language into a machine understandable representation

Vector Space Representations -

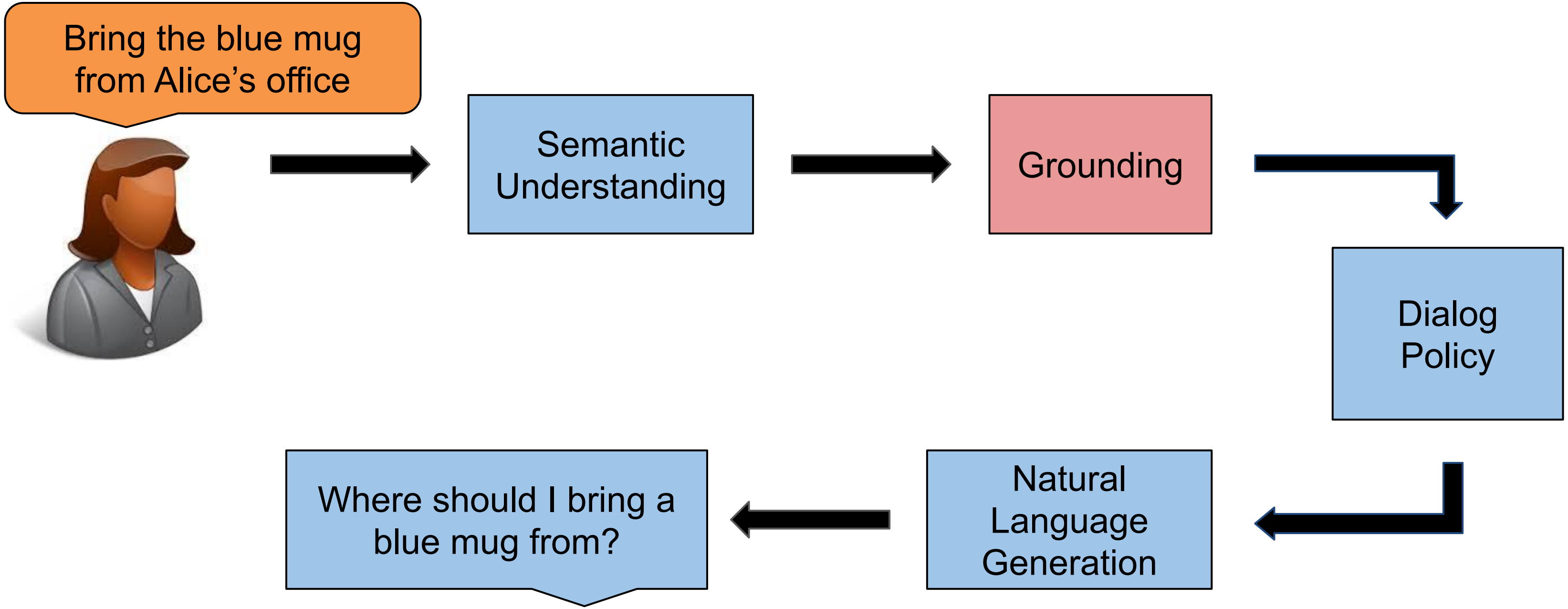
- Converts words/sentences to vectors that represent meaning.
- Less initial handcrafting
- More training data



Background: Grounding



Background: Grounding



Background: Grounding

Map meaning representations to real world entities

Background: Grounding

Map meaning representations to real world entities

$$\texttt{the}(\lambda y.(\texttt{office}(y) \wedge \texttt{owns}(\texttt{alice}, y)))$$

$\downarrow$

Knowledge Base Grounding

Person	Office
alice	3502
bob	3324

$\rightarrow$ `3502`

Background: Grounding

Map meaning representations to real world entities

Perceptual Grounding

$$\lambda x.(\texttt{blue}(x) \wedge \texttt{mug}(x))$$

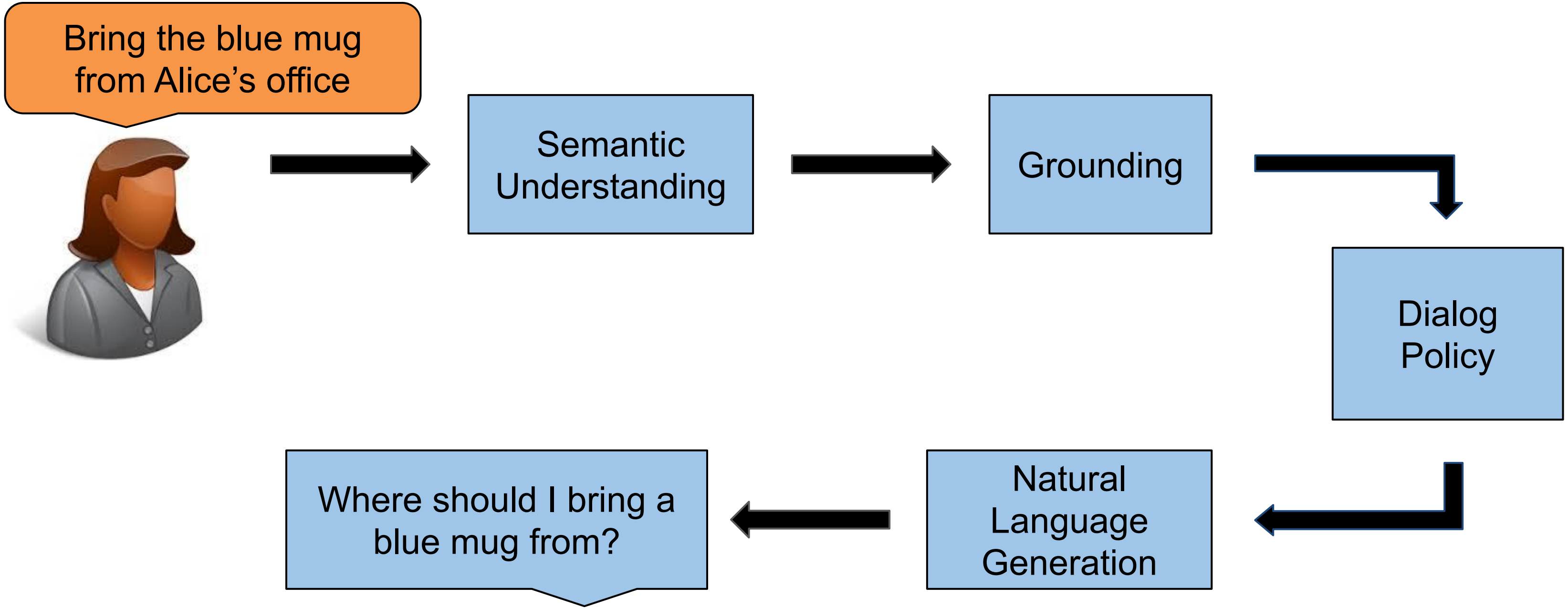
 → Classifier **`blue/not blue`** → **`blue`**

 → Classifier **`mug/not mug`** → **`mug`**

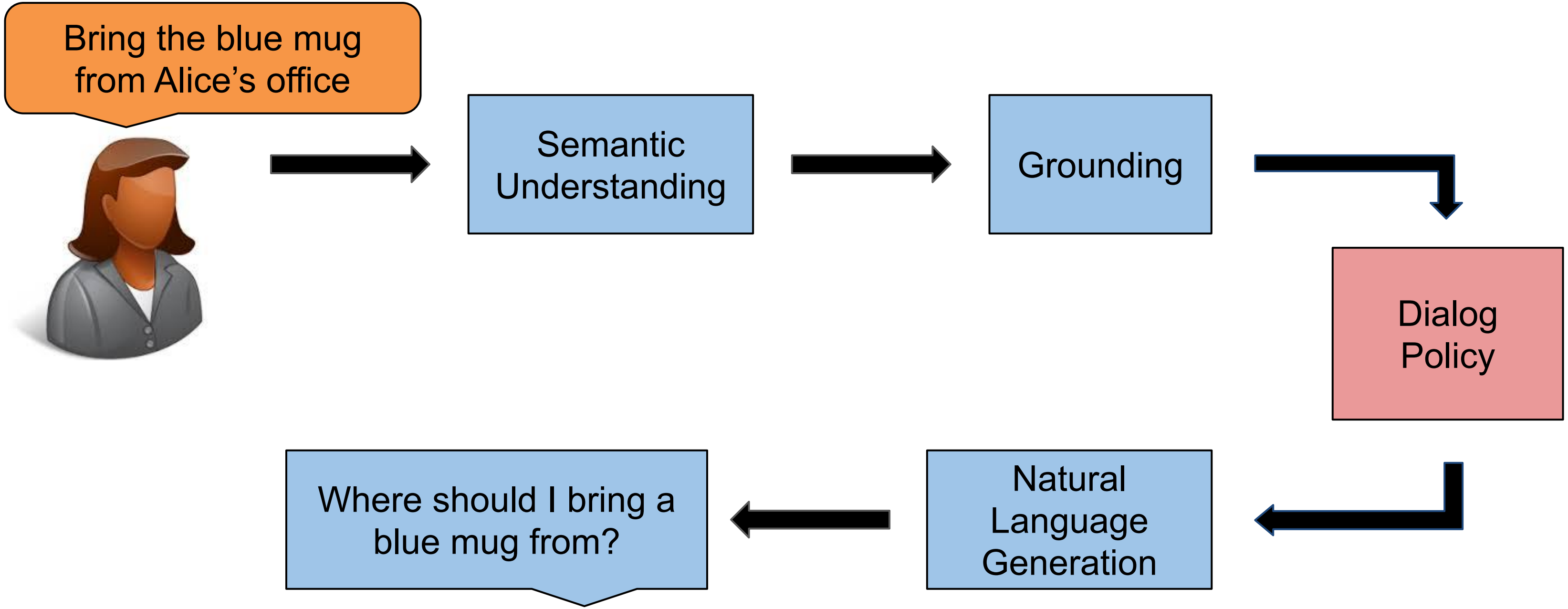
 → Classifier **`blue/not blue`** → **`not blue`**

 → Classifier **`mug/not mug`** → **`mug`**

Background: Dialog Policy



Background: Dialog Policy



Background: Dialog Policy

Plans the next response that the system has to give.

Bring the blue mug from Alice's office

Confirm

Ask Question

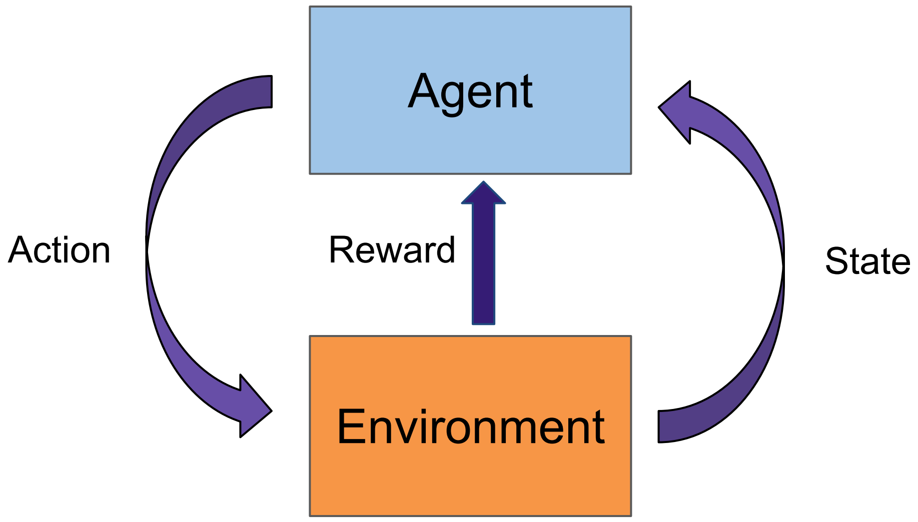
Execute

Background: Dialog Policy

- Dialog state - Information from the dialog so far
- Dialog policy - Mapping from dialog states to dialog actions (response types/ responses)
- Learned using Reinforcement Learning

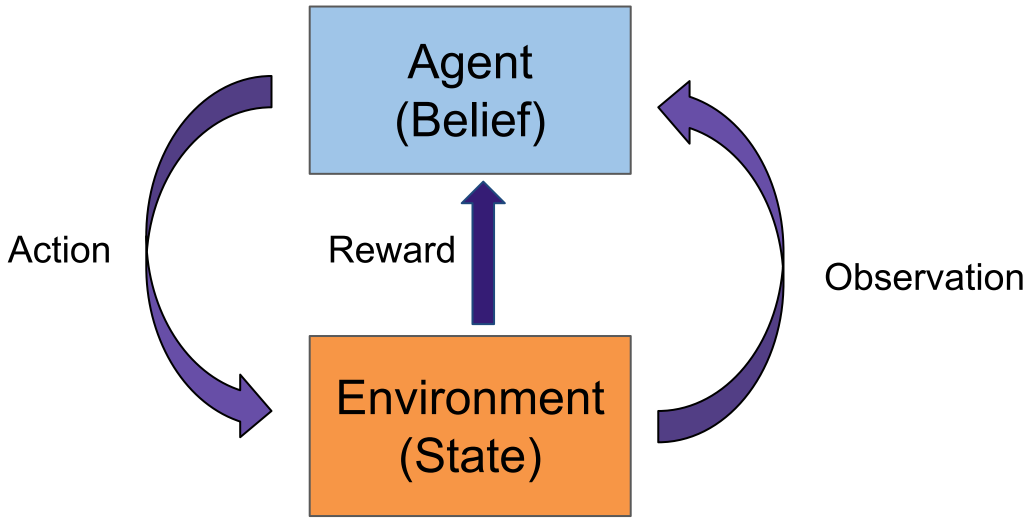
Background: Reinforcement Learning

Markov Decision Process (MDP)

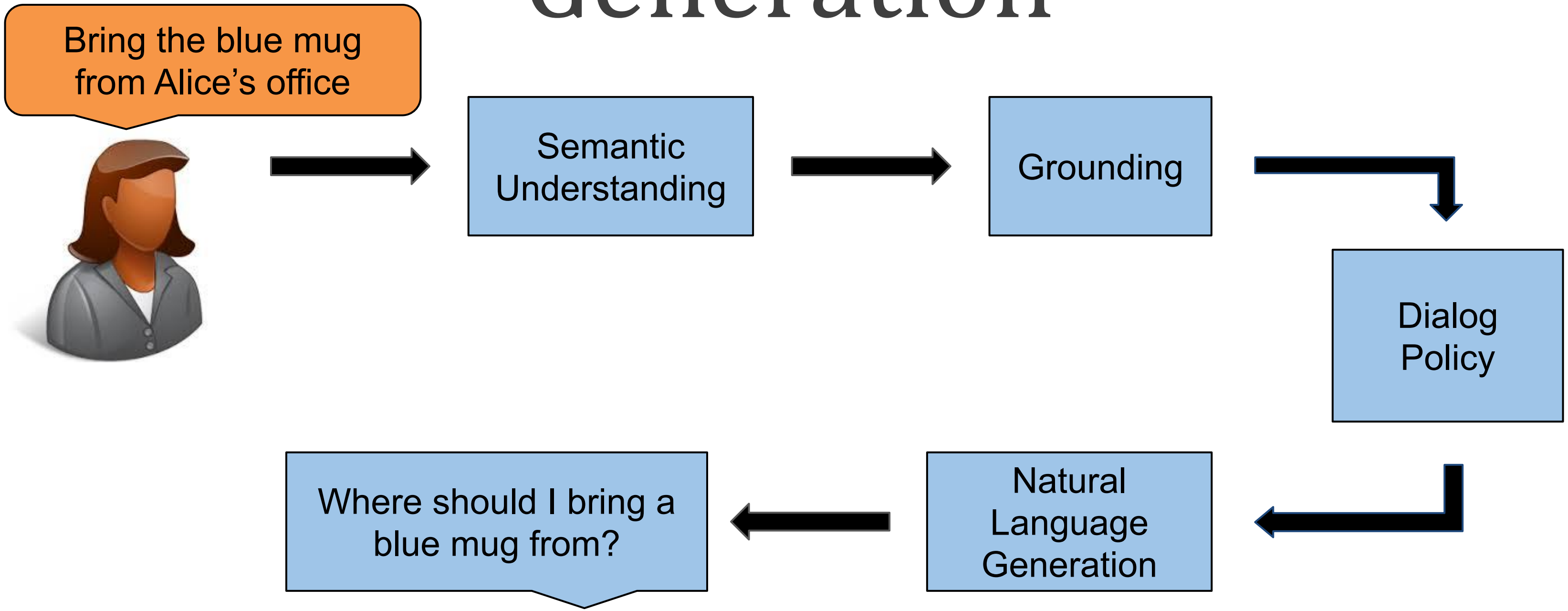


Background: Reinforcement Learning

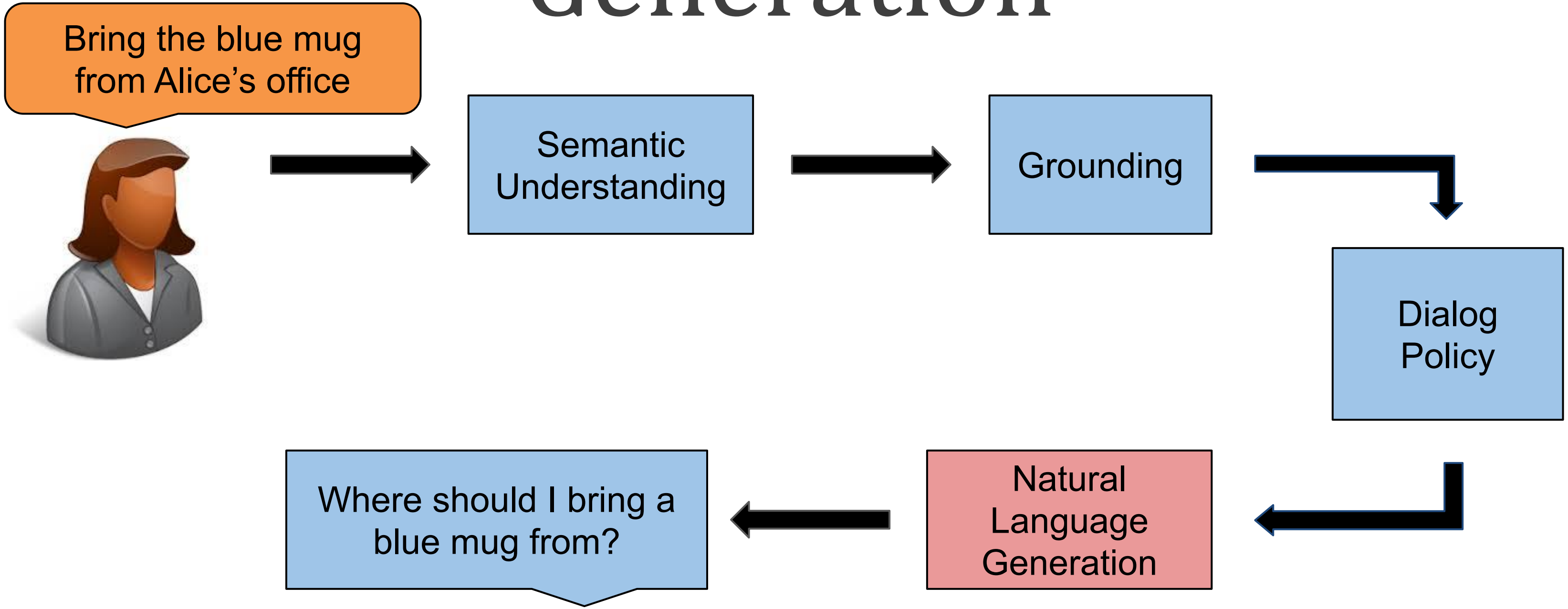
Partially Observable Markov Decision Process (POMDP)



Background: Natural Language Generation



Background: Natural Language Generation



Background: Natural Language Generation

Converting an action to a natural language response

```
ask_param(
    action=bring,
    patient=λx.(blue(x) ∧ mug(x))
    src=?
)
```

→

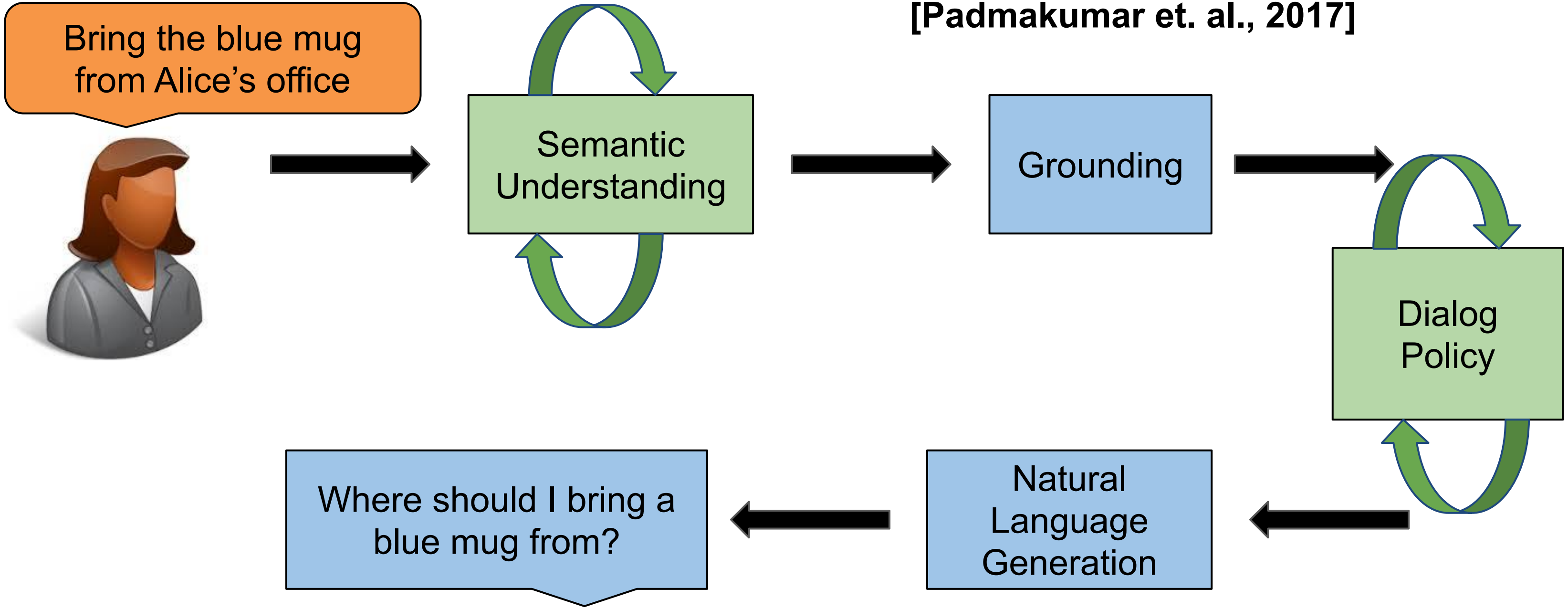
Where should I bring a blue mug from?

Outline

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Integrating Learning of Dialog Strategies and Semantic Parsing

[Padmakumar et. al., 2017]



Prior work: Improving Semantic Parsers from Clarification Dialogs

[Thomason et. al., 2015]

Bring the blue mug from Alice's office

Where should I bring a blue mug from?

Alice Ashcraft's office

I should bring a blue mug from 3502?

Yes

Alice's office

$\asymp$

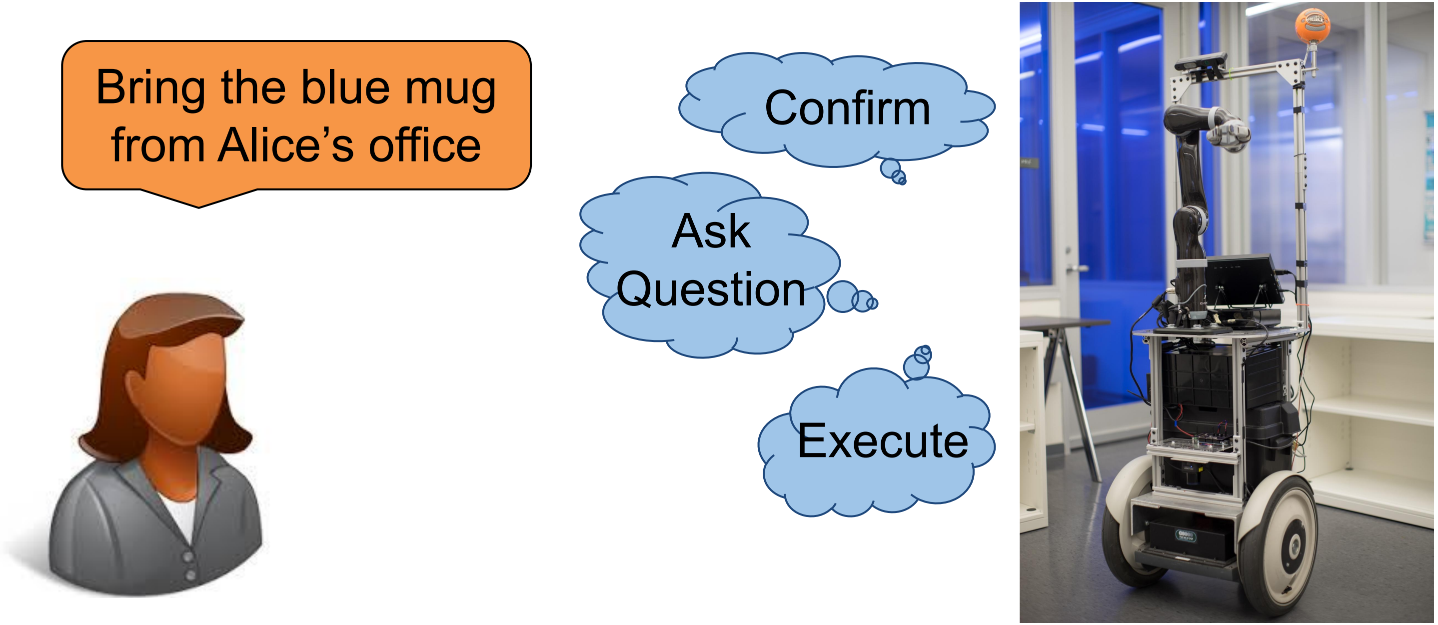
Alice Ashcraft's office

$\asymp$

3502

Prior Work: Dialog Policy Learning

Learns what the best next response is by modelling dialog system as a Partially Observable Markov Decision Process (POMDP)



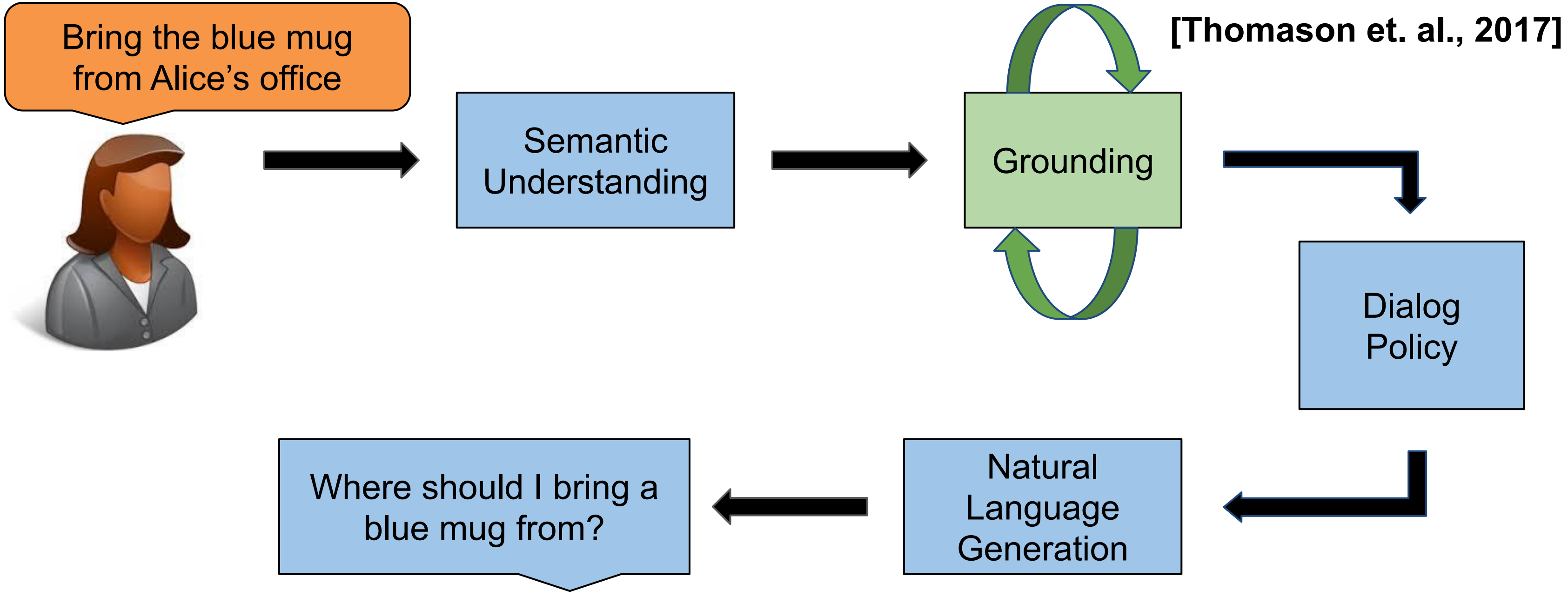
Summary

- Jointly improving a semantic parser and dialog policy from human interactions is more effective than improving either alone.
- The training procedure needs to enable changes in components to be propagated to each other for joint learning to be effective.

Outline

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Opportunistic Active Learning for Grounding Natural Language Descriptions



Opportunistic Active Learning

- A framework for incorporating active learning queries into test time interactions.
- Agent asks locally convenient questions during an interactive task to collect labeled examples for supervised learning.
- Questions may not be useful for the current interaction but expected to help future tasks.

Opportunistic Active Learning

Bring the blue mug
from Alice's office

Blue?

Opportunistic Active Learning

Bring the blue mug from Alice's office

Would you use the word "blue" to refer to this object?

Yes

Opportunistic Active Learning

Bring the blue mug from Alice's office

`bring(`  `,3502)`

Heavy?

Tall?

Opportunistic Active Learning

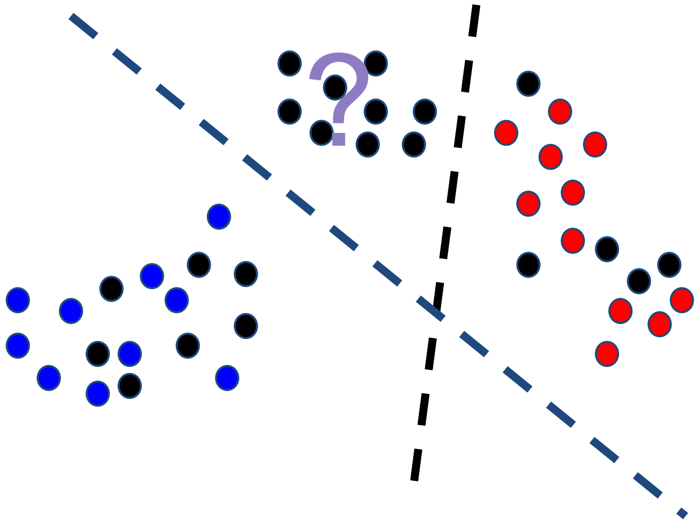
Bring the **blue mug** from Alice's office

Would you use the word "**tall**" to refer to this object?

Yes

Opportunistic Active Learning

Query for labels most likely to improve the model.



Opportunistic Active Learning

Why ask off-topic 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

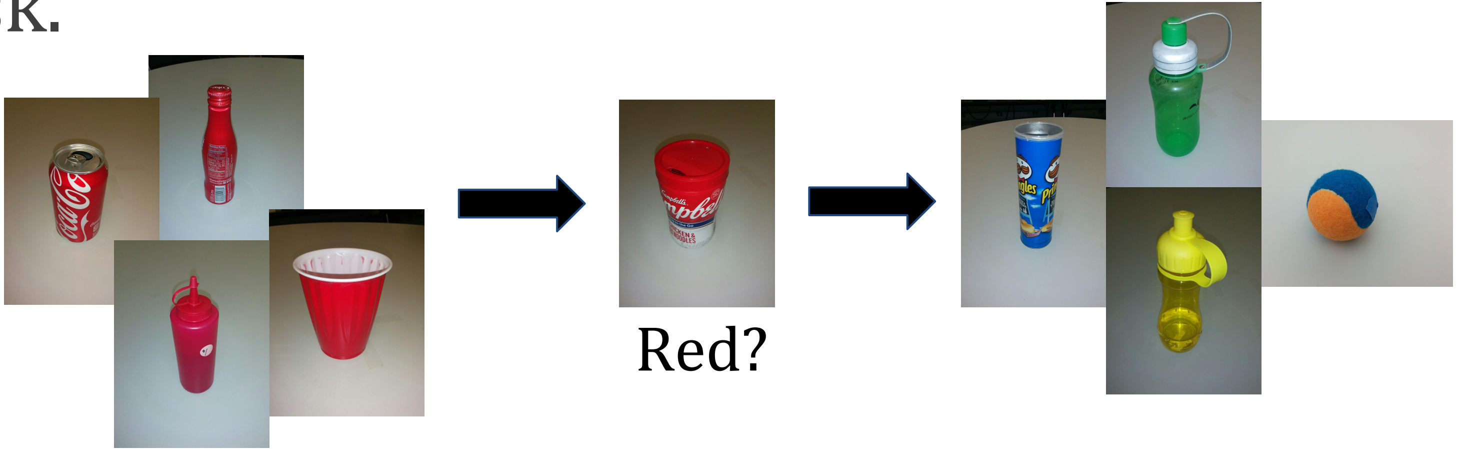
The question interrupts another task and may be seen as unnatural

Bring the **blue mug** from Alice's office

Would you use the word "**tall**" to refer to this object?

Opportunistic Active Learning - Challenges

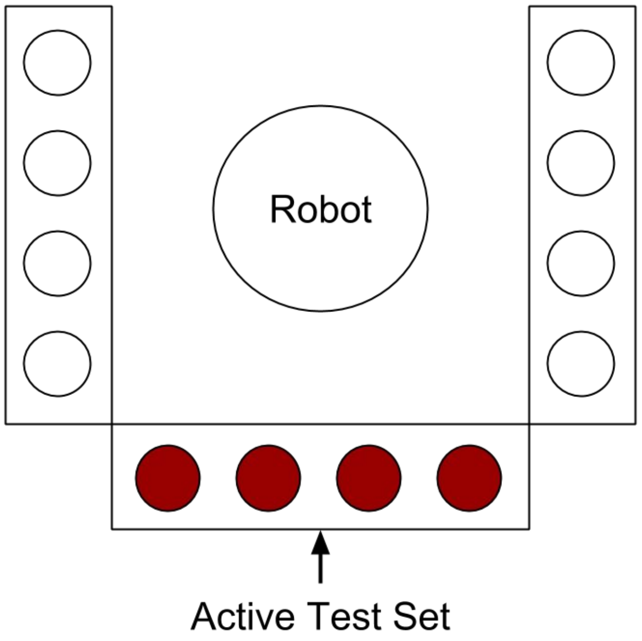
The information needs to be useful for a future task.



Object Retrieval Task

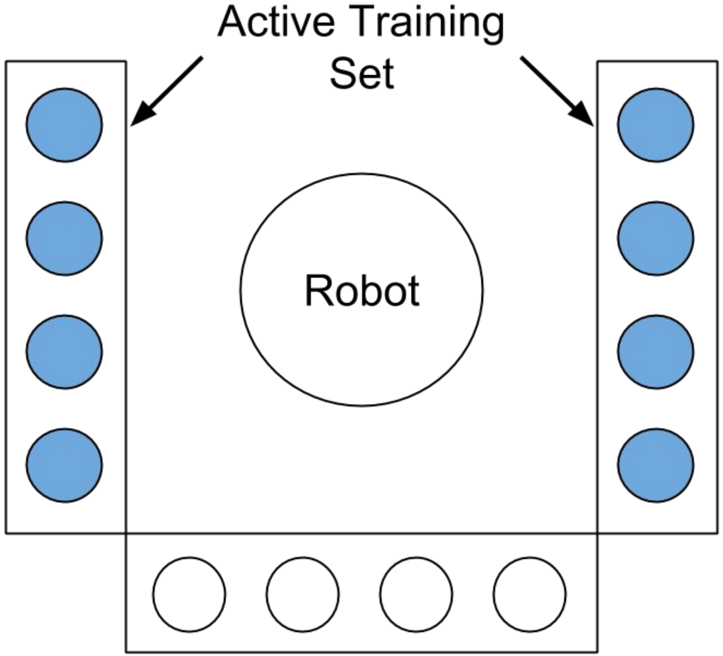
Object Retrieval Task

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

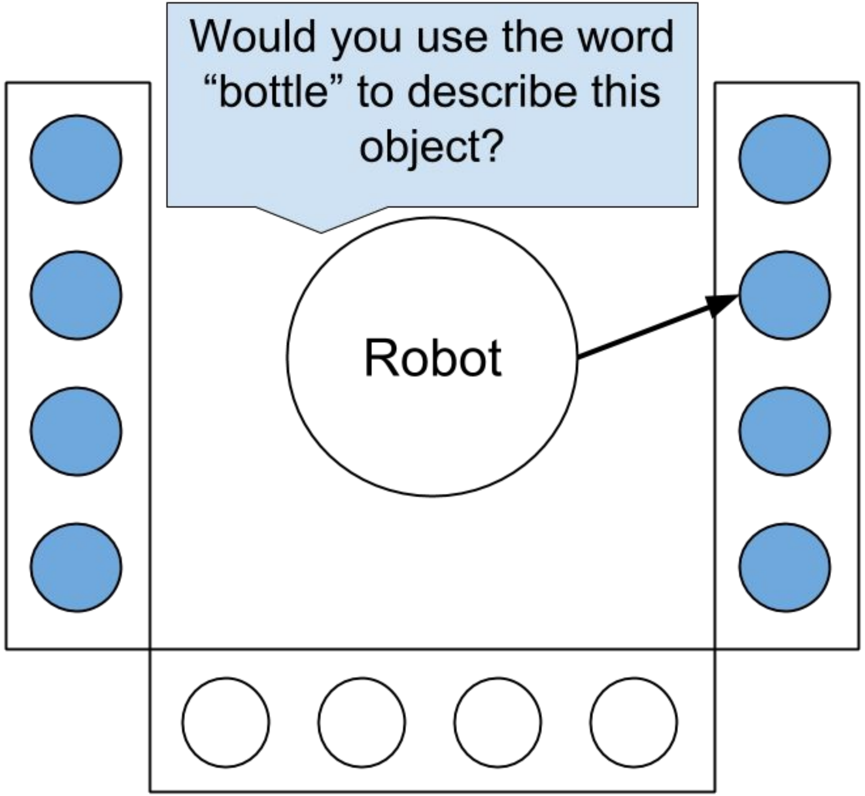


Object Retrieval Task

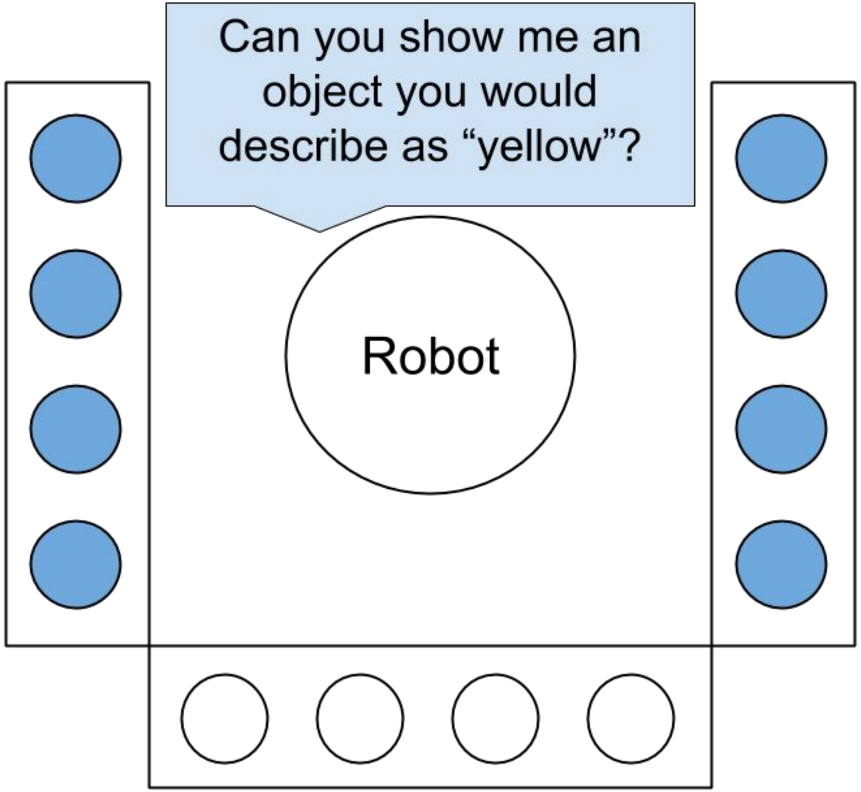
- Robot can ask questions about objects on the sides to learn object attributes



Two Types of Questions



Two Types of Questions



Experimental Conditions

A yellow water bottle

- Baseline (on-topic) - the robot can only ask about "yellow", "water" and "bottle"
- Inquisitive (on and off topic) - the robot can ask about any concept it knows, possibly "red" or "heavy"

Results

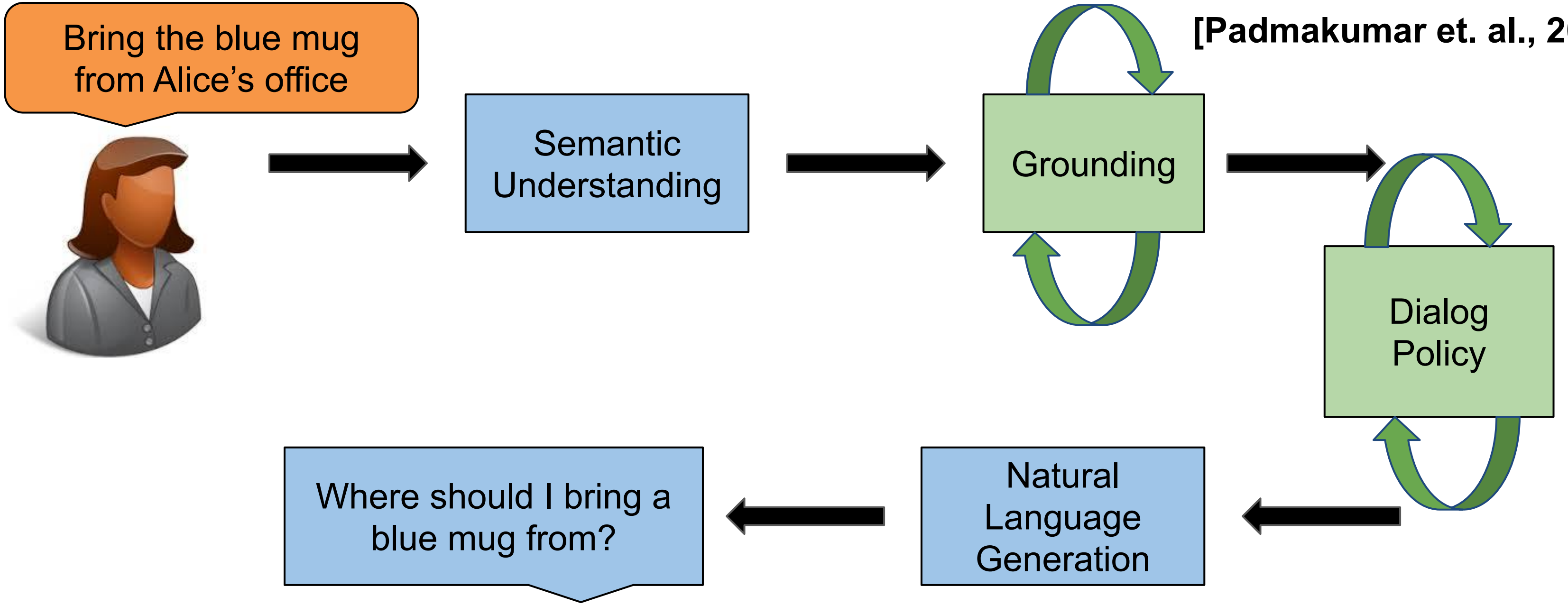
- Inquisitive robot performs better at understanding object descriptions.
- Users find the robot more comprehending, fun and usable in a real-world setting, when it is opportunistic.

Outline

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Learning a Policy for Opportunistic Active Learning

[Padmakumar et. al., 2018]



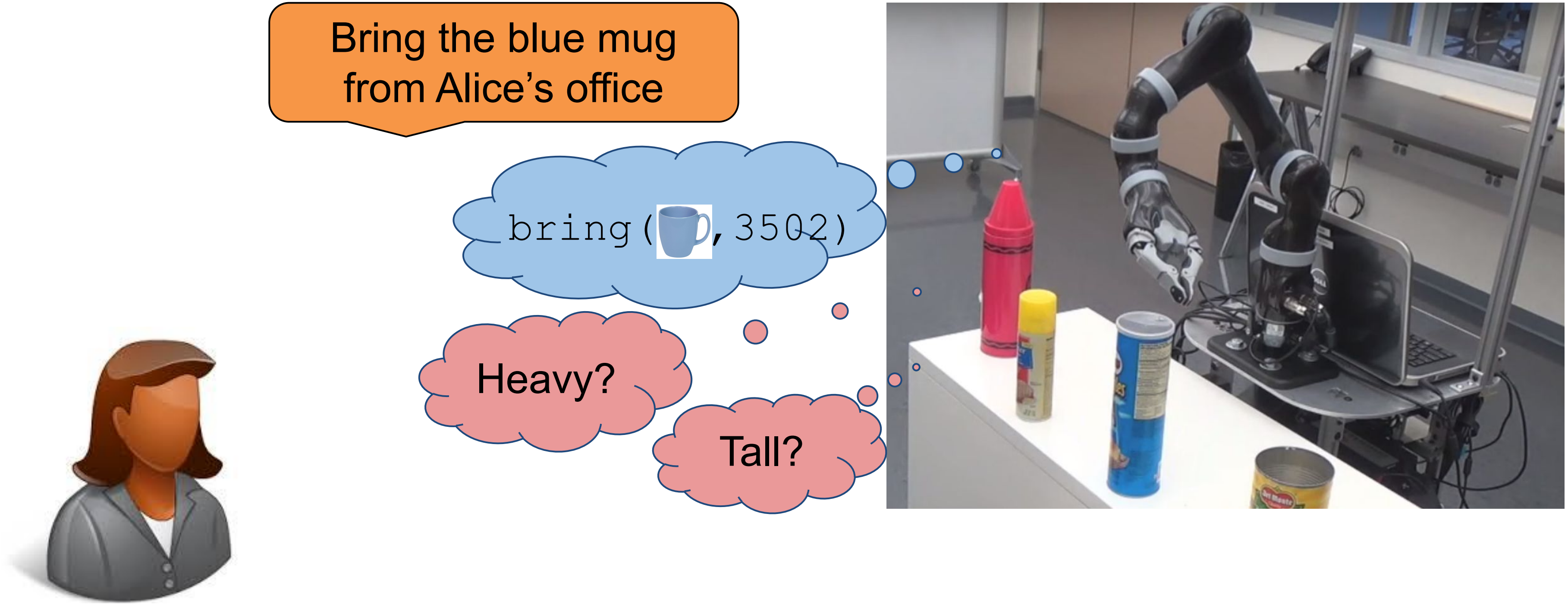
Opportunistic Active Learning

Bring the **blue mug** from Alice's office

Would you use the word "**tall**" to refer to this object?

Yes

Dialog Policy Learning



Learning a Policy for Opportunistic Active Learning

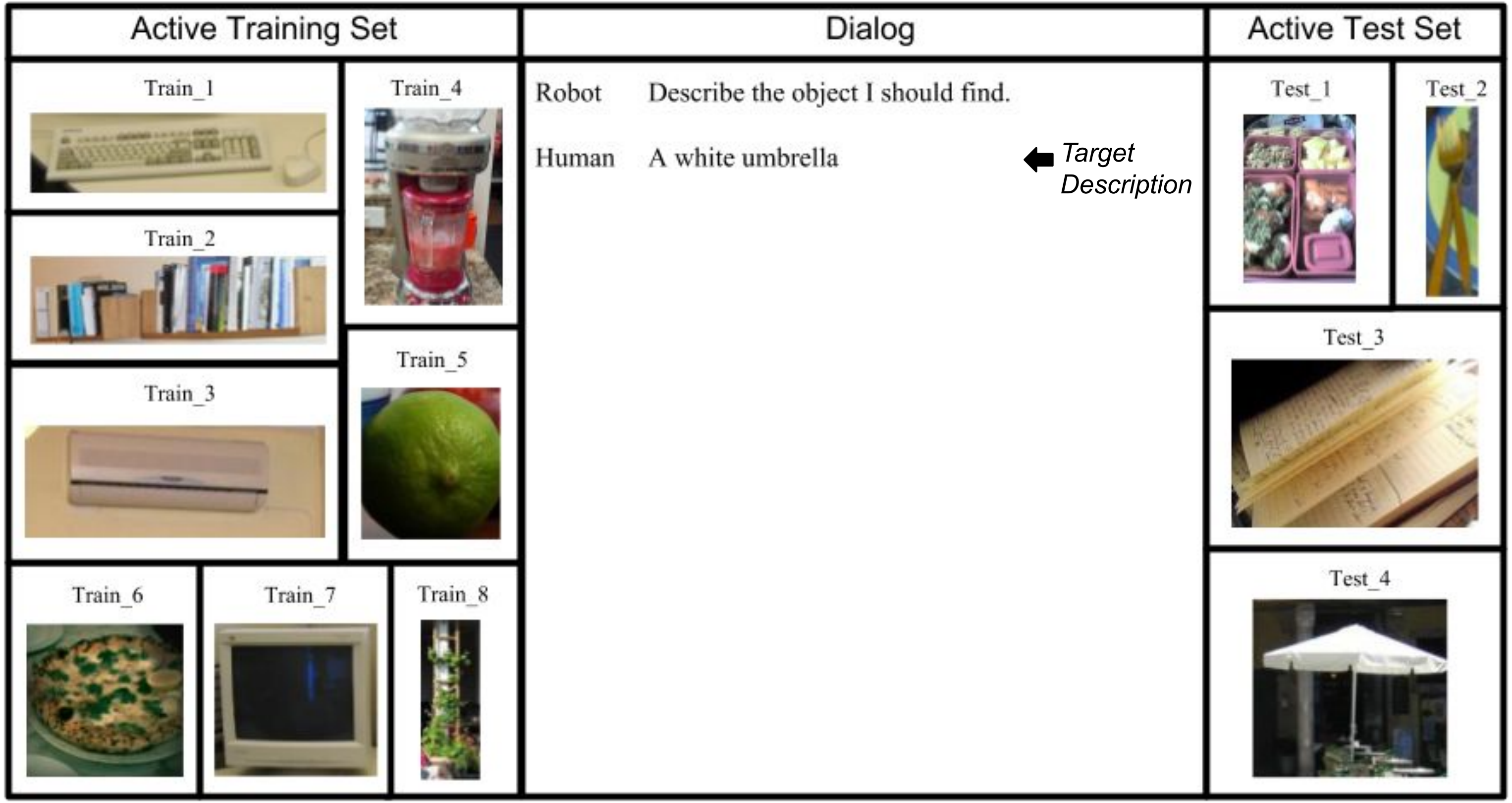
Learn a dialog policy that decides how many and which questions to ask to improve grounding models.

Learning a Policy for Opportunistic Active Learning

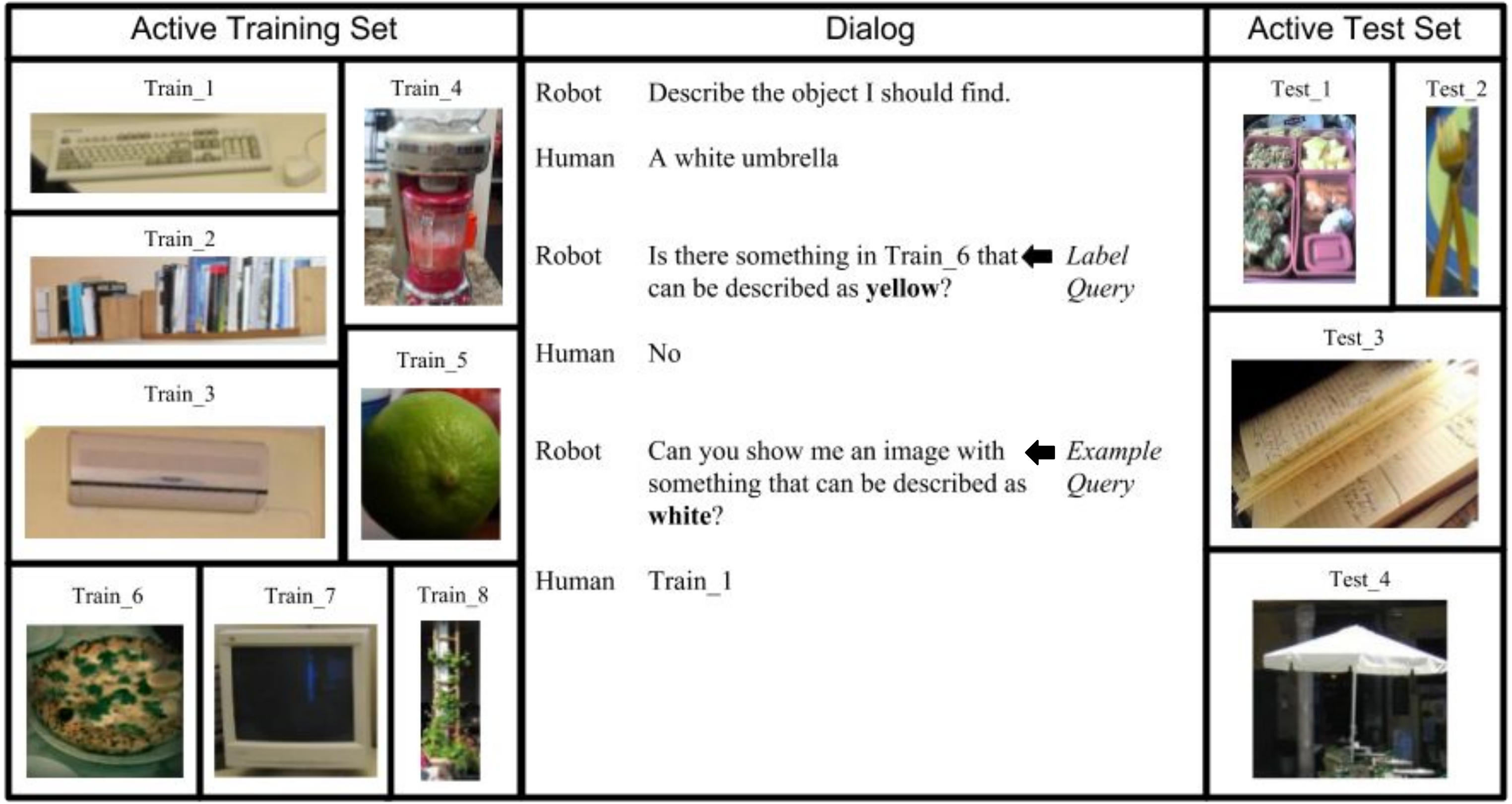
To learn an effective policy, the agent needs to learn

- To identify good queries in the opportunistic setting.
- When a guess is likely to be successful.
- To trade off between model improvement and task completion.

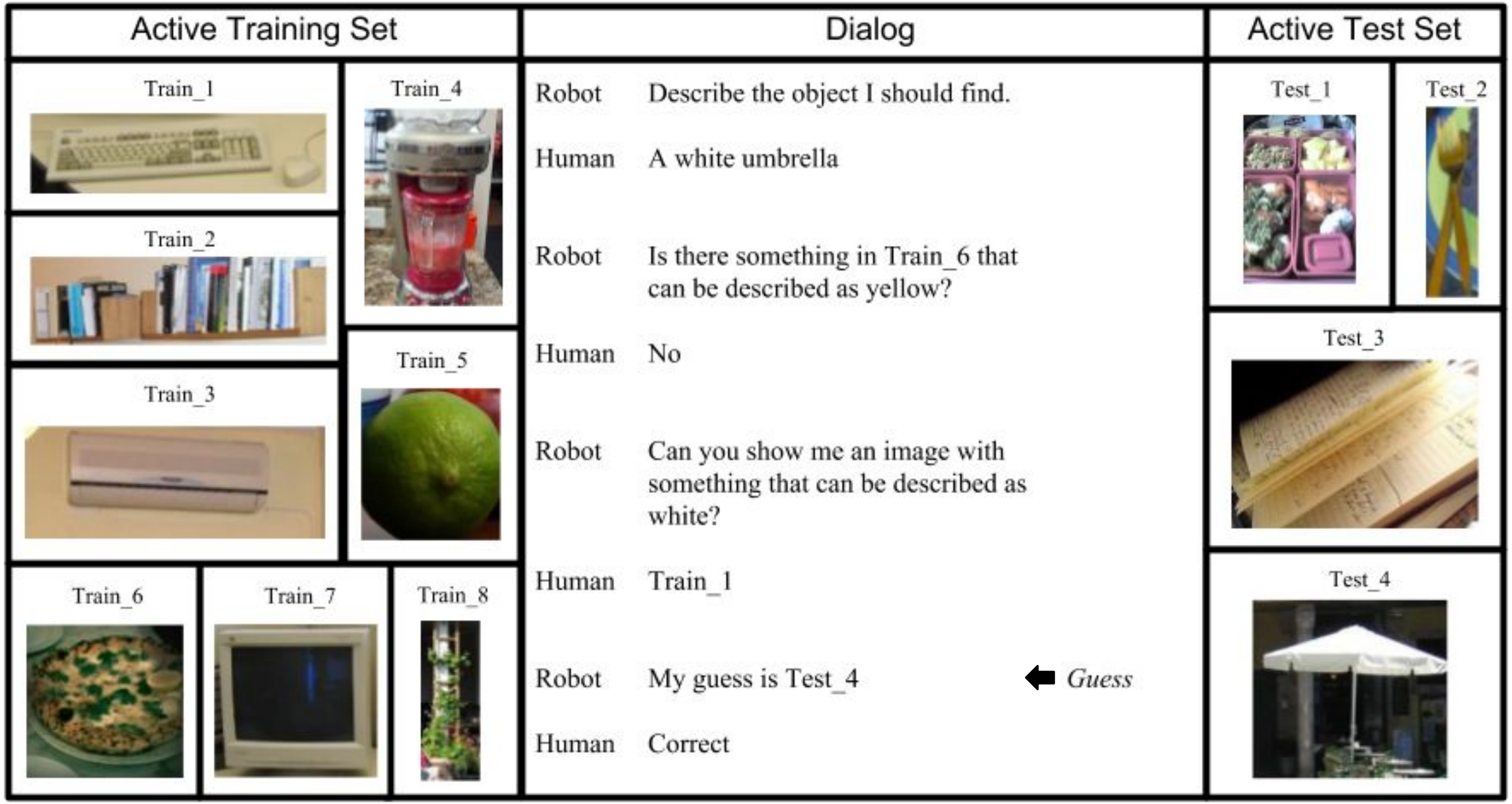
Task Setup


Active Training Set
Dialog
Active Test Set
Train_1
Train_2
Train_3
Train_4
Train_5
Train_6
Train_7
Train_8
Robot Describe the object I should find.
Human A white umbrella
Target Description
Test_1
Test_2
Test_3
Test_4

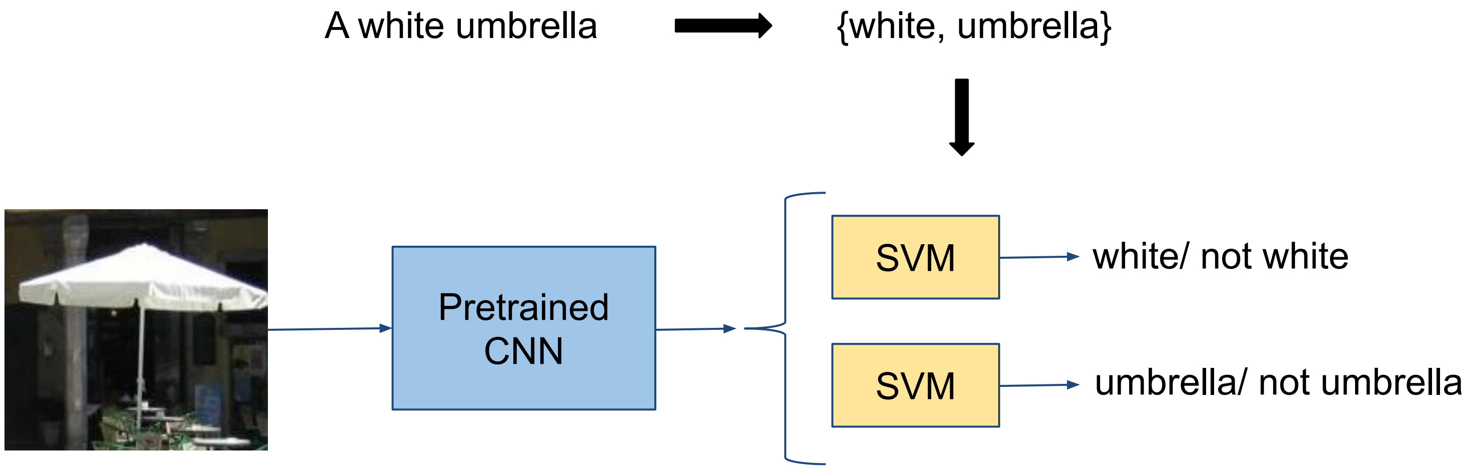
Task Setup


Active Training Set
Dialog
Active Test Set
Train_1
Train_2
Train_3
Train_4
Train_5
Train_6
Train_7
Train_8
Robot Describe the object I should find.
Human A white umbrella
Robot Is there something in Train_6 that can be described as **yellow**?
Label Query
Human No
Robot Can you show me an image with something that can be described as **white**?
Example Query
Human Train_1
Test_1
Test_2
Test_3
Test_4

Task Setup


Active Training Set
Dialog
Active Test Set
Train_1
Train_4
Train_2
Train_5
Train_3
Train_6
Train_7
Train_8
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
Robot My guess is Test_4
Guess
Human Correct
Test_1
Test_2
Test_3
Test_4

Grounding Model

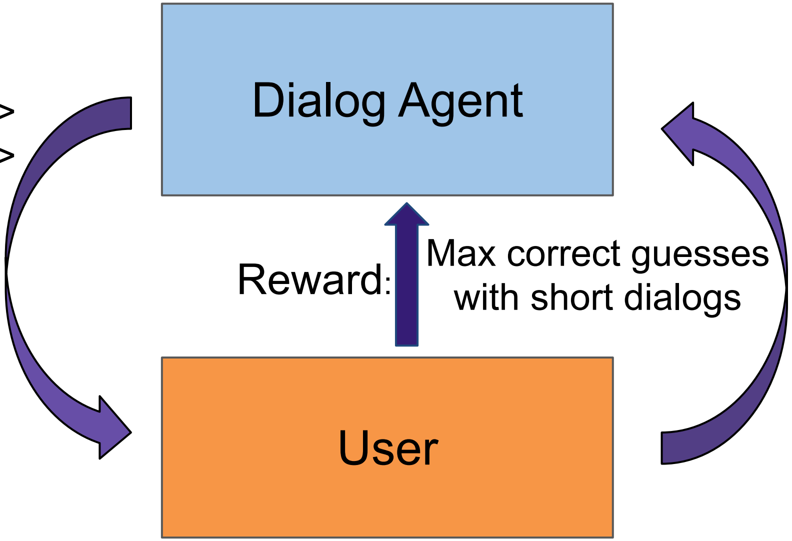


Opportunistic Active Learning

- 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.

MDP Model

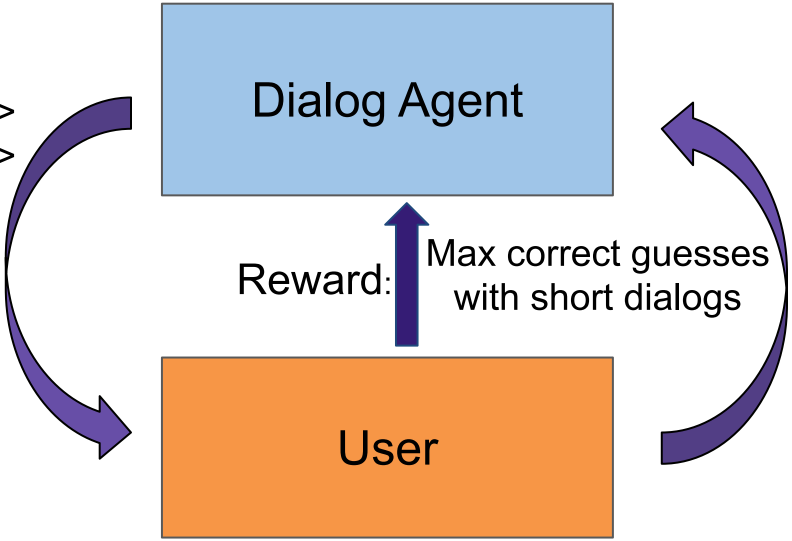
Action:
- Label query: <yellow, train_1>
- Label query: <yellow, train_2>
- …
- Label query: <white, train_1>
- Label query: <white, train_2>
- ...
- Example Query: yellow
- Example query: white
- ...
- Guess



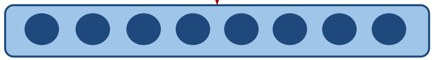
State:
- Target description
- Active train and test objects
- Agent's perceptual classifiers

Challenges

Action:
- Label query: <yellow, train_1>
- Label query: <yellow, train_2>
- …
- Label query: <white, train_1>
- Label query: <white, train_2>
- ...
- Example Query: yellow
- Example query: white
- ...
- Guess



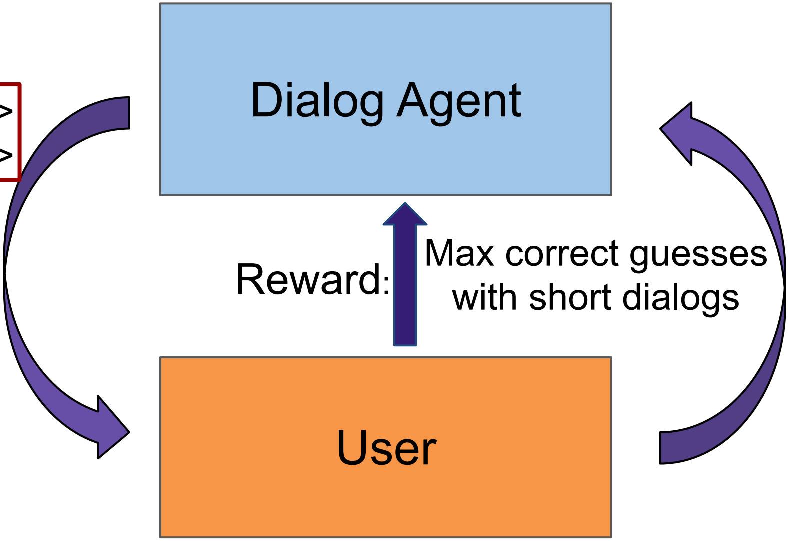
State:
- Target description
- Active train and test objects
- Agent’s perceptual classifiers



How to represent classifiers for policy learning?

Challenges

Action:
- Label query: <yellow, train_1>
- Label query: <yellow, train_2>
- …
- Label query: <white, train_1>
- Label query: <white, train_2>
- ...
- Example Query: yellow
- Example query: white
- ...
- Guess



State:
- Target description
- Active train and test objects
- Agent's perceptual classifiers

How to handle a variable and growing action space?

Tackling challenges

- Features based on active learning metrics
  - Representing classifiers
- Featurize state-action pairs
  - Variable number of actions and classifiers
- Sampling a beam of promising queries
  - Large action space

Feature Groups

- Query features - Active learning metrics used to determine whether a query is useful
- Guess features - Features that use the predictions and confidences of classifiers to determine whether a guess will be correct

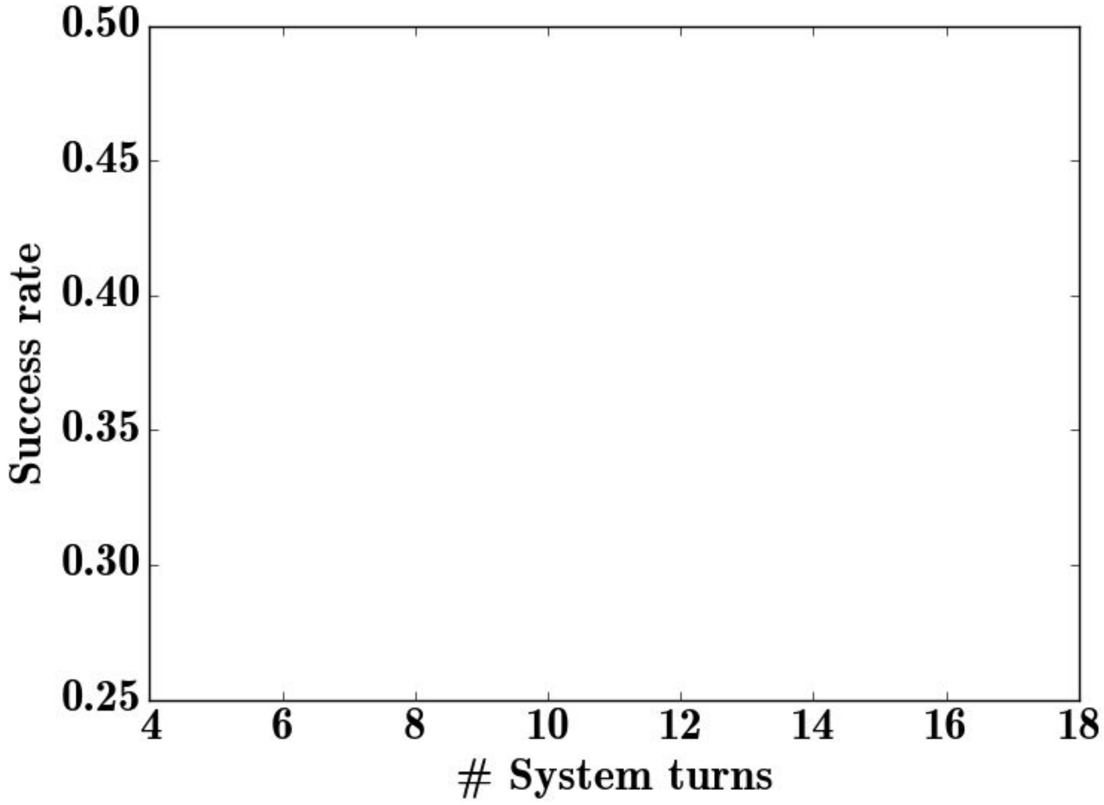
Experiment Setup

- Policy learning using REINFORCE.
- Baseline - A hand-coded dialog policy that asks a fixed number of questions selected using the sampling distribution that provides candidates to the learned policy.

Experiment Phases

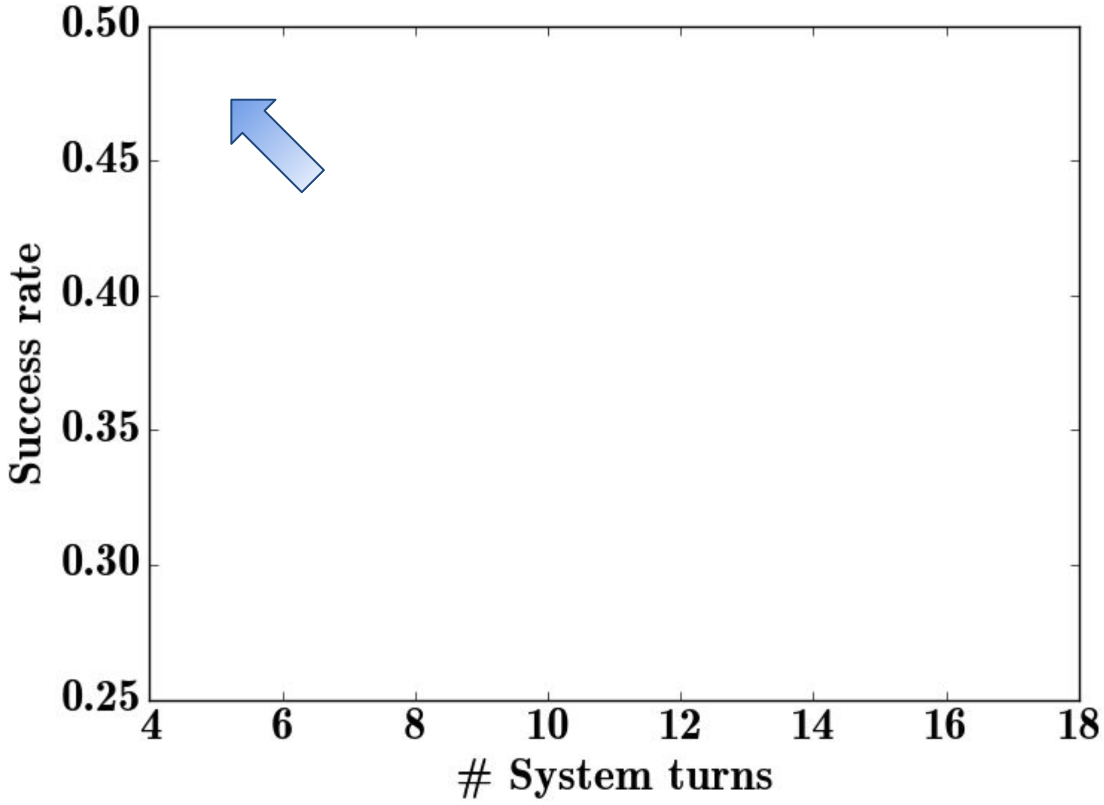
- Initialization - Collect experience using the baseline to initialize the policy.
- Training - Improve the policy from on-policy experience.
- Testing - Policy weights are fixed, and we run a new set of interactions, starting with no classifiers, over an independent test set with different predicates.

Results



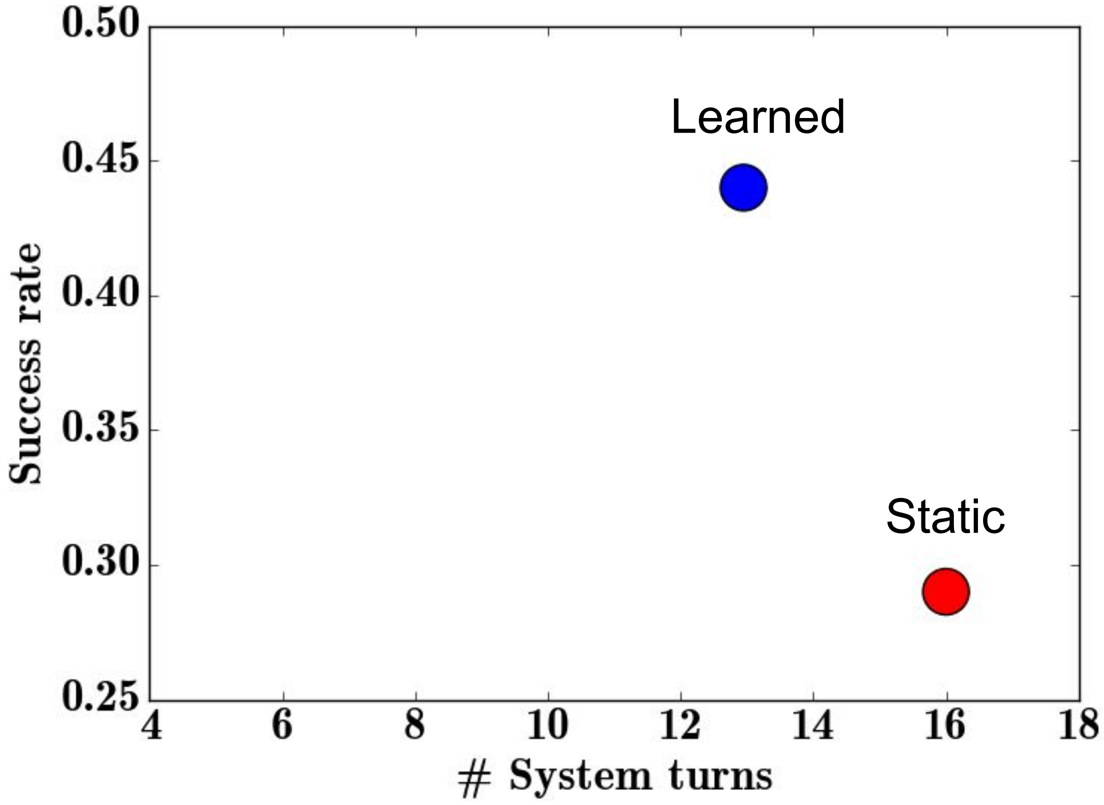
- Systems evaluated on dialog success rate and average dialog length.

Results



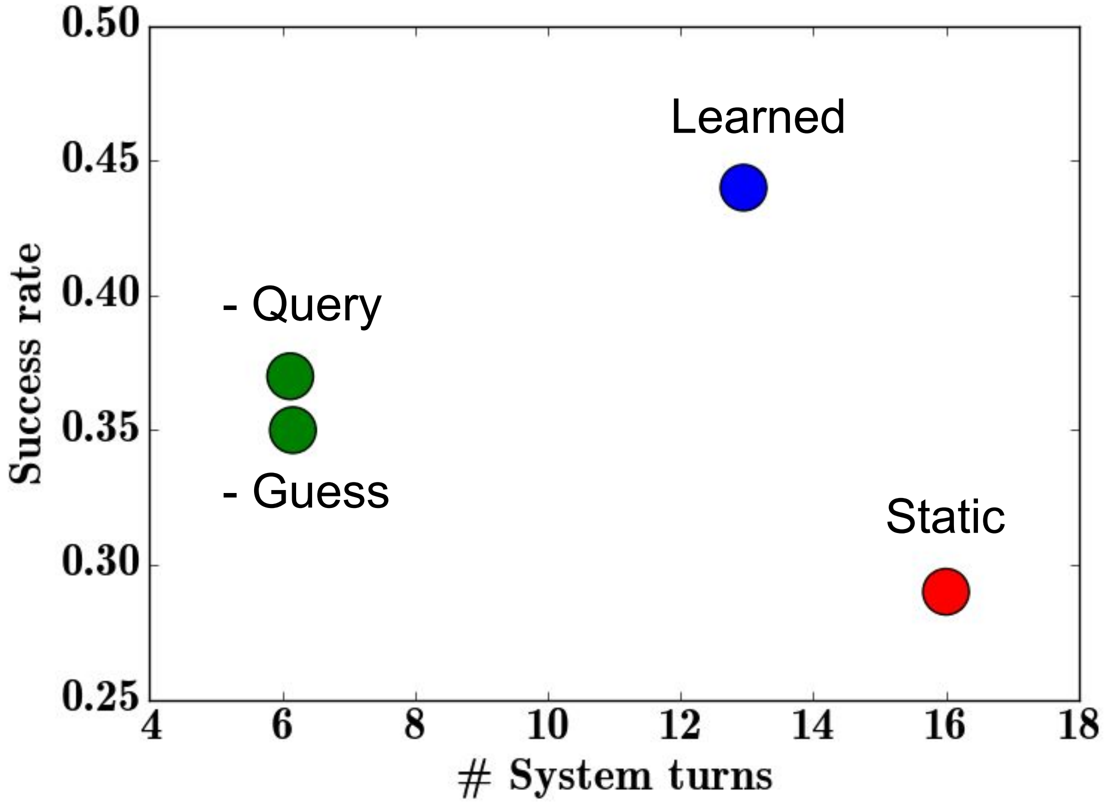
- Systems evaluated on dialog success rate and average dialog length.
- We prefer high success rate and low dialog length (top left corner)

Results



- Learned policy is more successful than the baseline, while also using shorter dialogs on average.

Results



- 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.

Outline

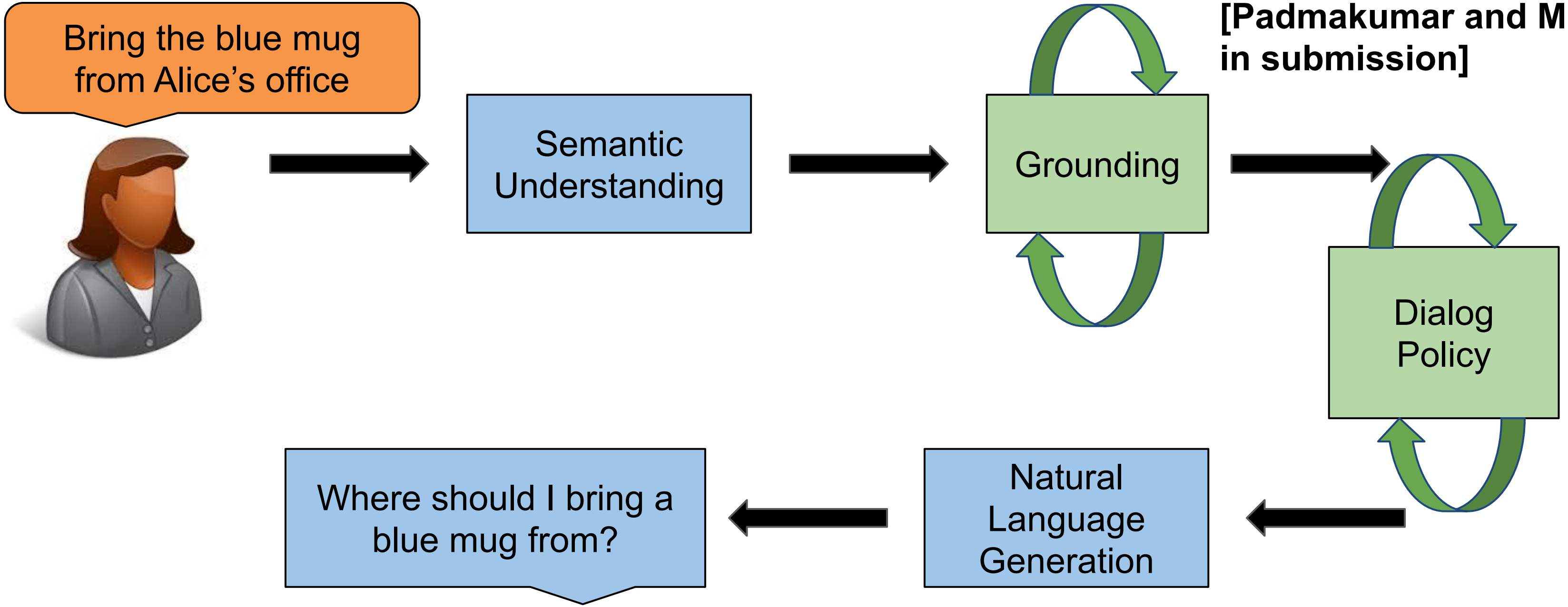
- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Outline

- Dialog Policy Learning for Joint Clarification and Active Learning Queries
  - Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
  - Human Evaluation
  - Extension to Joint Embedding Based Grounding Model

Dialog Policy Learning for Joint Clarification and Active Learning Queries

[Padmakumar and Mooney, in submission]



Previous Work

Bring the blue mug from Alice's office

`bring(``,3502)`

Heavy?

Tall?

This Work

Bring the blue mug from Alice's office

`bring(●,3502)`

Heavy?

Tall?

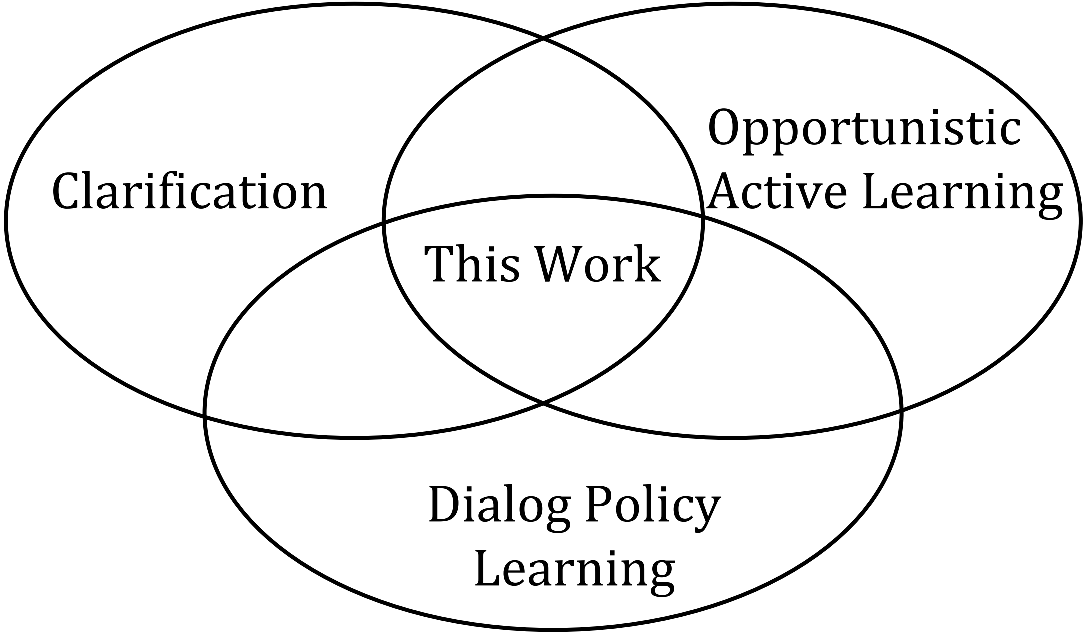
This Work

Bring the blue mug from Alice's office

What should I bring?

Would you use the word "tall" to refer to this object?

Dialog Policy Learning for Joint Clarification and Active Learning Queries



Dialog Policy Learning for Joint Clarification and Active Learning Queries

Learn a dialog policy to trade off -

- Model improvement with opportunistic active learning to better understand future commands
- Clarification to better understand and complete the current command

Attribute Based Clarification: Motivation

Bring the blue mug from Alice's office

bring(●, 3502)

What should I bring?

Attribute Based Clarification: Motivation

Bring the blue mug from Alice's office

What should I bring?

The blue coffee mug

What should I bring?

Attribute Based Clarification: Motivation

Bring the blue mug from Alice's office

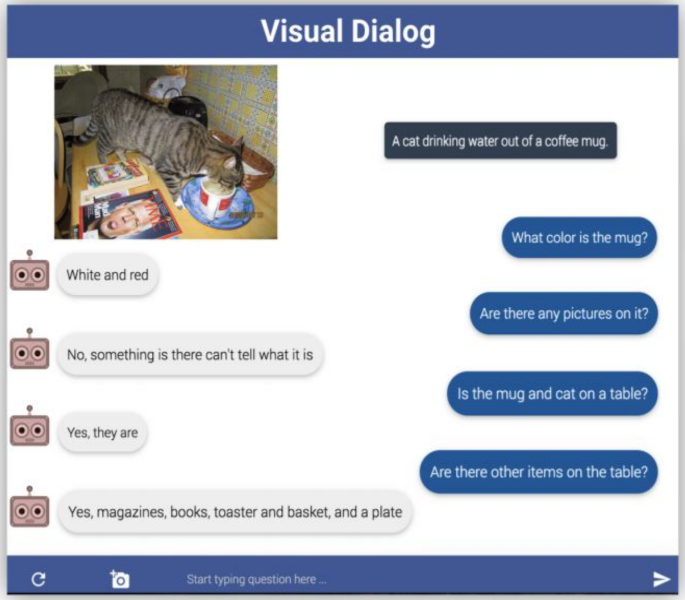
Is this the object I should bring?

No

Is this the object I should bring?

Attribute Based Clarification: Motivation

[Das, et. al., 2017]



[De Vries et. al., 2017]

Questioner	Oracle
Is it a vase?	Yes
Is it partially visible?	No
Is it in the left corner?	No
Is it the turquoise and purple one?	Yes

Attribute Based Clarification

- More specific than a new description.
- More general than showing each possible object.
- Provide ground truth answers to questions for training in simulation.
- Attribute - any property that can be used in a description - categories, colors, shapes, domain specific properties.

Attribute Based Clarification: Motivation

Bring the blue mug from Alice's office

Is the object I should bring a cup?

Task Setup

- Motivated by an online shopping application
- Use clarifications to help refine search queries
- Use active learning to improve the model retrieving images.

Dataset

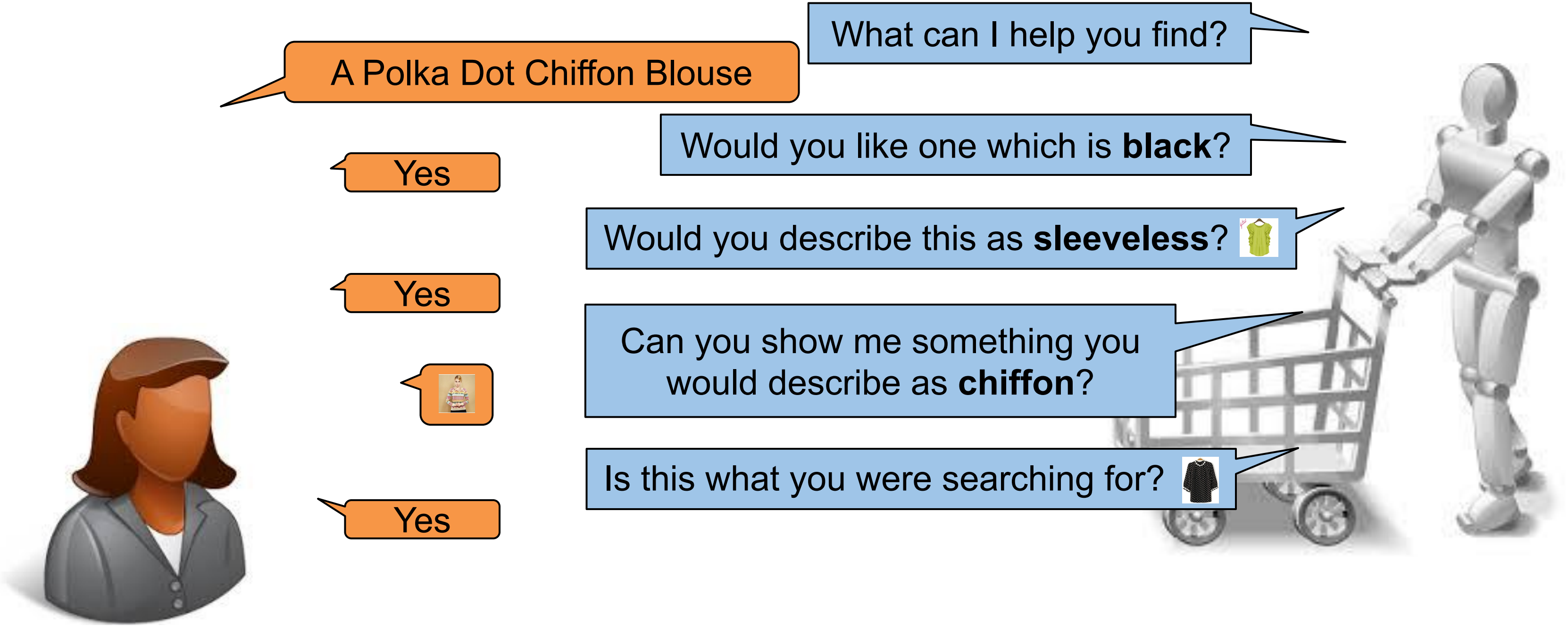
- We simulate dialogs using the iMaterialist Fashion Attribute dataset.
- Images have associated product titles and are annotated with binary labels for 228 attributes.
- Attributes: Dress, Shirt, Red, Blue, V-Neck, Pleats, ...

Task Setup

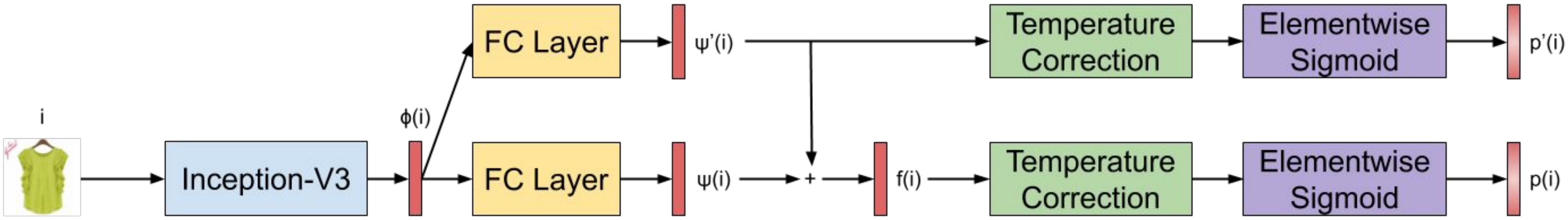
Active Training Set

Active Test Set

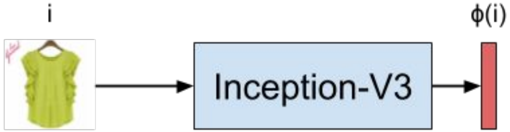
Task Setup


What can I help you find?
A Polka Dot Chiffon Blouse
Would you like one which is **black**?
Yes
Would you describe this as **sleeveless**?
Yes
Can you show me something you would describe as **chiffon**?
Is this what you were searching for?
Yes

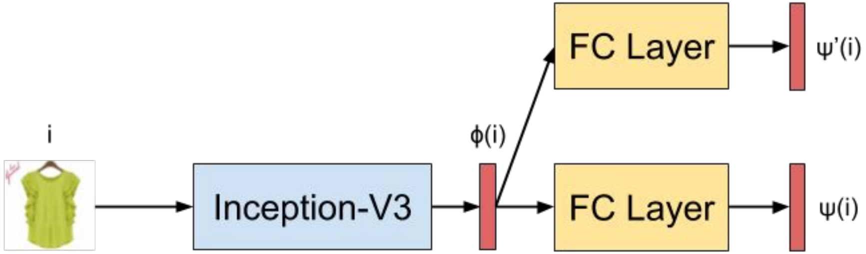
Visual Attribute Classifier



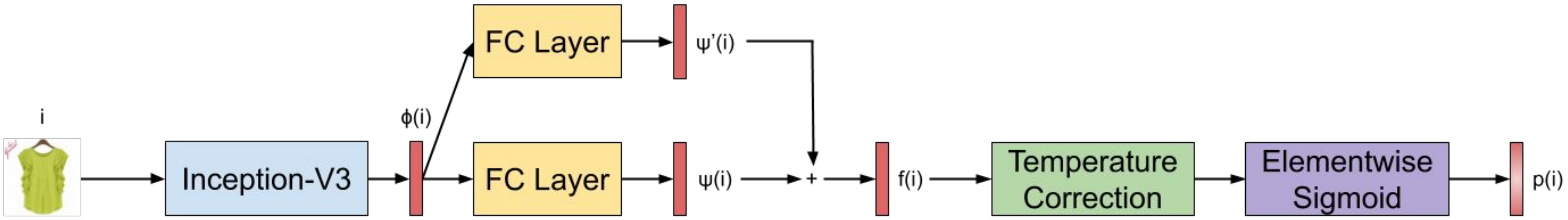
Visual Attribute Classifier



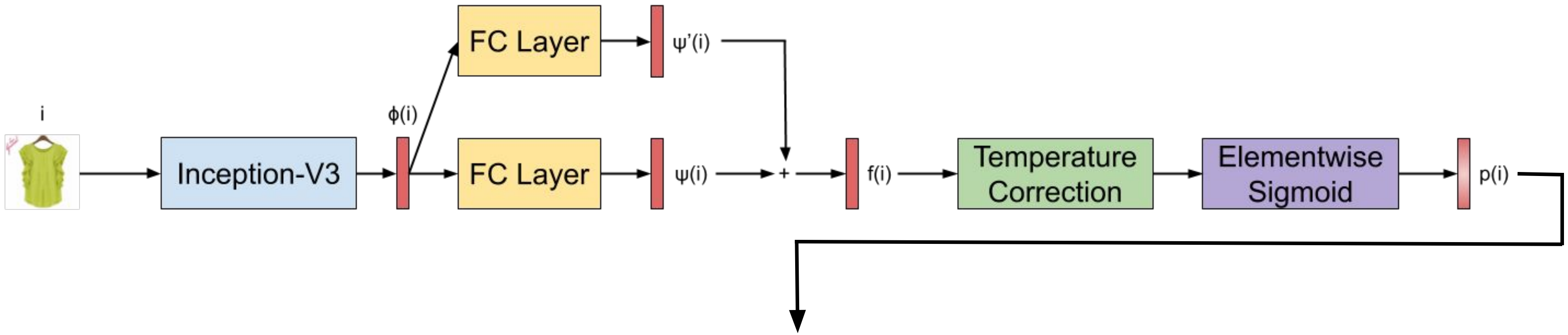
Visual Attribute Classifier



Visual Attribute Classifier



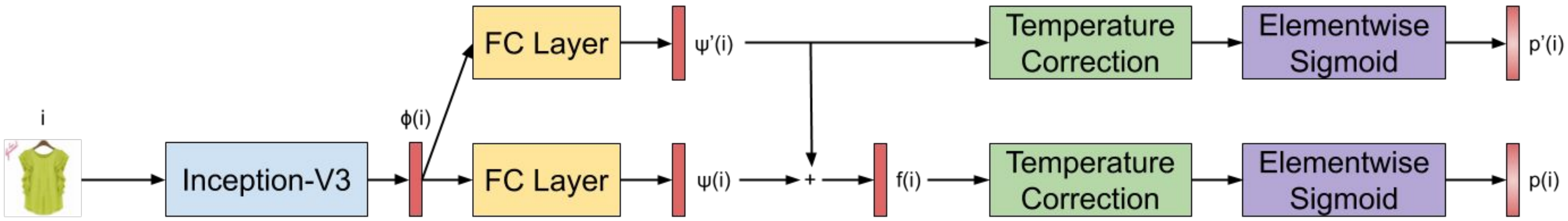
Visual Attribute Classifier



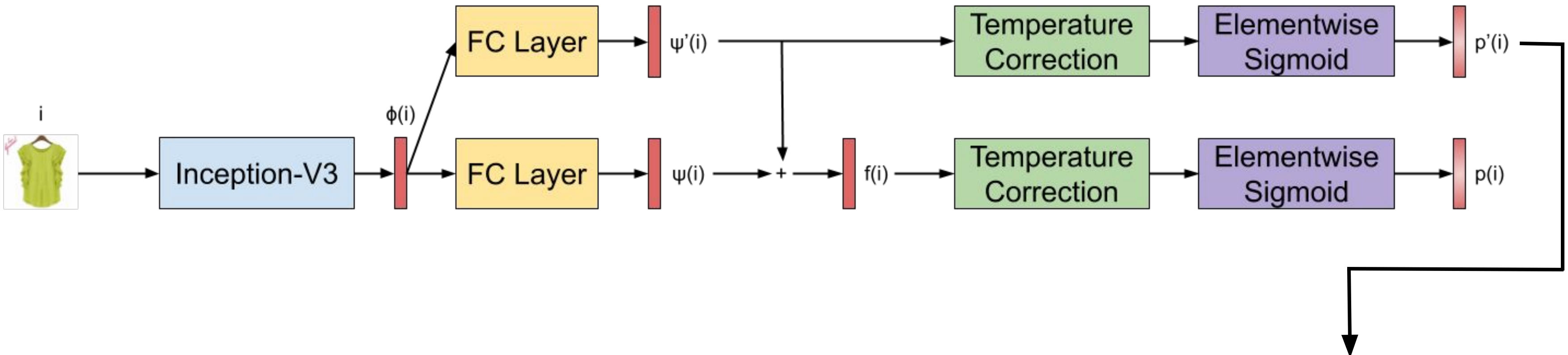
$$L = (1-\lambda)\boxed{\sum_i y_i \log p(i) + (1-y_i)\log(1-p(i))} + \lambda \sum_i y_i \log p'(i)$$

Cross Entropy Loss Over All Examples

Visual Attribute Classifier



Visual Attribute Classifier



$$L = (1-\lambda)\sum_i y_i \log p(i) + (1-y_i)\log(1-p(i)) + \lambda \boxed{\sum_i y_i \log p'(i)}$$

Cross Entropy Loss Over Positive Labels

Grounding Model

A Polka Dot Chiffon Blouse $\longrightarrow$ {Polka Dot, Chiffon, Blouse}

Grounding Model

A Polka Dot Chiffon Blouse → {Polka Dot, Chiffon, Blouse}

Belief: $b(i) = \prod_{\boxed{w \in W_d}} p_w(i)$

Attributes Mentioned in Description

Grounding Model

A Polka Dot Chiffon Blouse $\longrightarrow$ {Polka Dot, Chiffon, Blouse}

Belief: $b(i) = \prod_{w \in W_d} \boxed{p_w(i)}$

- Classifier probability that attribute w is positive for image i
- `w`-th value in classifier output for image `i`

Grounding Model

Agent: Would you like one which is black? → <Black, 1>

User: Yes

Belief: $$b(i) = \prod_{w \in W_d} p_w(i) \prod_{\boxed{w \in W_p}} p_w(i)$$

Clarifications that get the answer "Yes"

Grounding Model

Agent: Would you like one which is black?

<Black, 0>

User: No

Belief: $$b(i) = \prod_{w \in W_d} p_w(i) \prod_{w \in W_p} p_w(i) \prod_{\boxed{w \in W_n}} (1 - p_w(i))$$

Clarifications that get the answer “No”

Grounding Model

Best guess: Image in active test set with maximum belief b(i)

Information Gain

- For estimating the utility of clarifications
- Estimated using classifier probabilities
- Estimate based on Lee et. al., 2018

Information Gain

$$J(q) = \sum_{i \in O_A^{te}} \sum_{a \in \{0,1\}} b(i) P(a|q,i) \ln\left(\frac{P(a|q,i)}{\sum_i b(i) P(a|q,i)}\right)$$

Information Gain

$$J(q) = \sum_{\boxed{i \in O_A^{te}}} \sum_{a \in \{0,1\}} b(i) P(a|q,i) \ln \left(\frac{P(a|q,i)}{\sum_i b(i) P(a|q,i)} \right)$$

Objects in Active Test Set

Information Gain

$$J(q) = \sum_{i \in O_A^{te}} \sum_{\boxed{a \in \{0,1\}}} b(i) P(a|q,i) \ln \left(\frac{P(a|q,i)}{\sum_i b(i) P(a|q,i)} \right)$$

Possible answers to a clarification: No and Yes

Information Gain

$$J(q) = \sum_{i \in O_A^{te}} \sum_{a \in \{0,1\}} \boxed{b(i)} P(a|q,i) \ln \left(\frac{P(a|q,i)}{\sum_i \boxed{b(i)} P(a|q,i)} \right)$$

Belief of image i

Information Gain

$$J(q) = \sum_{i \in O_A^{te}} \sum_{a \in \{0,1\}} b(i) \boxed{P(a|q,i)} \ln \left(\frac{\boxed{P(a|q,i)}}{\sum_i b(i) \boxed{P(a|q,i)}} \right)$$

Probability of the answer

Information Gain

$$J(q) = \sum_{i \in O_A^{te}} \sum_{a \in \{0,1\}} b(i) \boxed{P(a|q,i)} \ln \left(\frac{\boxed{P(a|q,i)}}{\sum_i b(i) \boxed{P(a|q,i)}} \right)$$

Probability of the answer

For “Yes” Answer: $P(1|q,i) = p_q(i)$

Information Gain

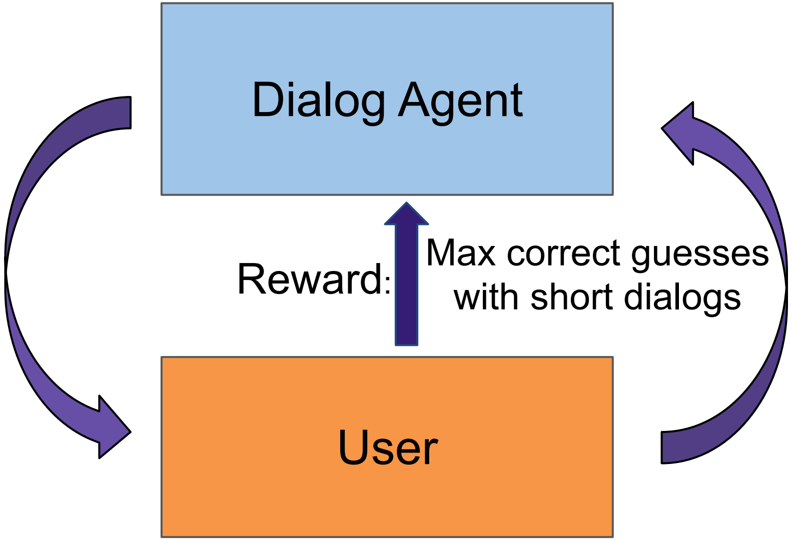
$$J(q) = \sum_{i \in O_A^{te}} \sum_{a \in \{0,1\}} b(i) \boxed{P(a|q,i)} \ln \left(\frac{\boxed{P(a|q,i)}}{\sum_i b(i) \boxed{P(a|q,i)}} \right)$$

Probability of the answer

For “No” Answer: $P(0|q,i) = 1 - p_q(i)$

Dialog as MDP

Action:
- Clarifications
- Label queries
- Example queries
- Guess



State:
- Target description
- Active train and test objects
- Agent's perceptual classifiers

Policy Learning

- Hierarchical Dialog Policy -
  - Clarification policy - chooses best clarification
  - Active learning policy - chooses best active learning query
  - Decision Policy - chooses between guess, best clarification and best active learning query
- Featurize state-action pairs
- Q-Learning and A3C for policy learning

Policy Features

- Clarification Policy Features - Metrics about current beliefs, information gain
- Active Learning Policy Features - Margin, Fraction of previous uses and successes
- Decision Policy Features - Metrics about current beliefs, information gain, margin, dialog length

Static Baseline

- Clarification: Choose query with maximum information gain
- Active Learning: Uncertainty Sampling
- Decision Policy
  - Fixed dialog length
  - Clarification till the belief reaches a threshold
  - Active learning for the second half of the dialog

Experiment Phases

- Classifier Initialization - Train classifier using paired images and labels
- Policy Initialization - Collect experience using the baseline to initialize the policy.
- Policy Training - Improve the policy from on-policy experience.
- Policy Testing - Policy weights are fixed, and we run a new set of interactions, reset classifiers to the state at the end of classifier initialization, over an independent test set with different predicates.

Results

Decision Policy Type	Clarification Policy Type	Active Learning Policy Type	Fraction of Successful Dialogs	Average Dialog Length
Q-Learning	A3C	A3C	**0.33**	9.40
Static	Static	Static	0.17	20.00

Results

Decision Policy Type	Clarification Policy Type	Active Learning Policy Type	Fraction of Successful Dialogs	Average Dialog Length
Q-Learning	A3C	A3C	**0.33**	9.40
Static	Static	Static	0.17	20.00

Fully learned policy is significantly more successful than the baseline, while also having significantly shorter dialogs on average

Results

Decision Policy Type	Clarification Policy Type	Active Learning Policy Type	Fraction of Successful Dialogs	Average Dialog Length
Q-Learning	A3C	A3C	**0.33**	9.40
Q-Learning	A3C	Static	0.15	14.16
Q-Learning	Static	A3C	0.09	1.00
Static	Static	Static	0.17	20.00

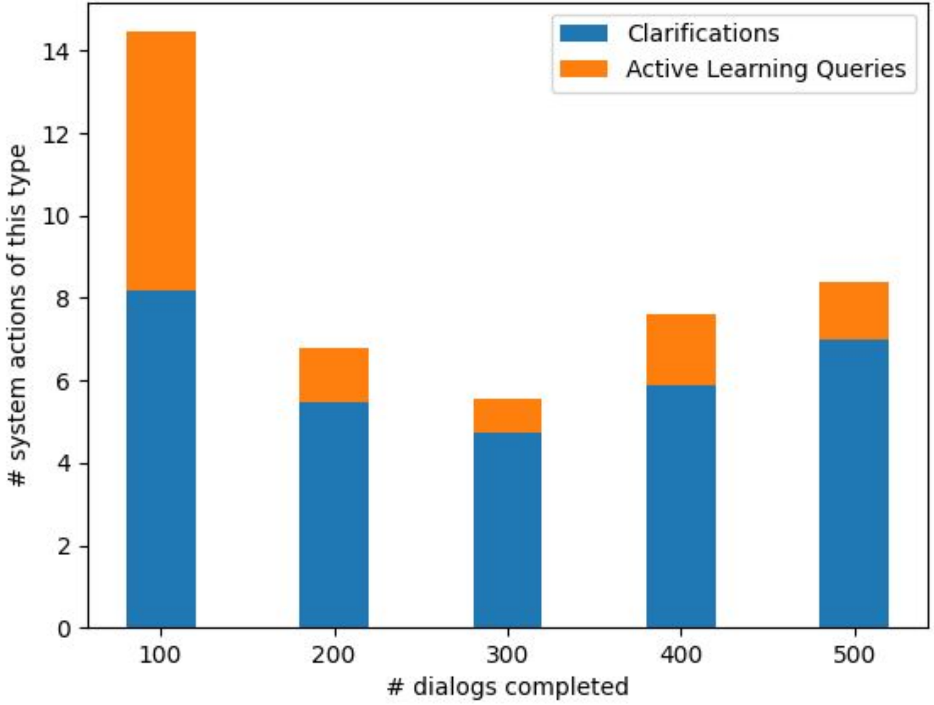
If we replace either the clarification or active learning policies with static policies, we find that the success rate drops considerably.

Results

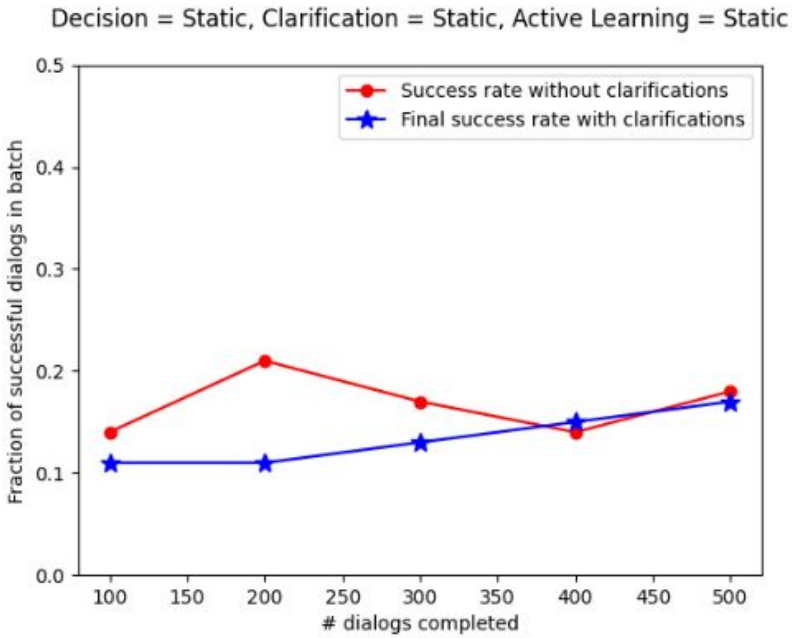
Decision Policy Type	Clarification Policy Type	Active Learning Policy Type	Fraction of Successful Dialogs	Average Dialog Length
Q-Learning	A3C	A3C	**0.33**	9.40
Static	A3C	A3C	0.27	20.00
Static	Static	Static	0.17	20.00

If we replace only the decision policy with a static policy, we find that it remains more successful than the baseline but is unable to shorten dialogs.

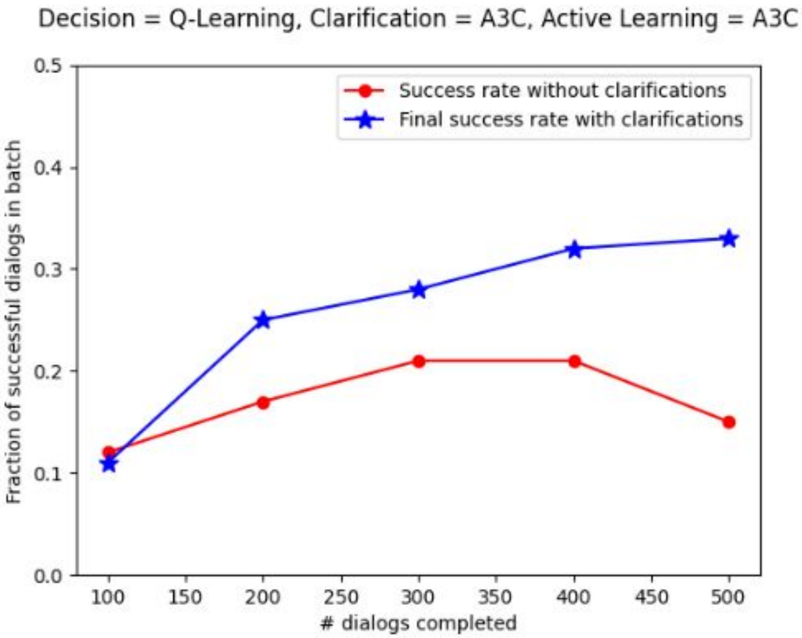
Action Types - Learned Policy

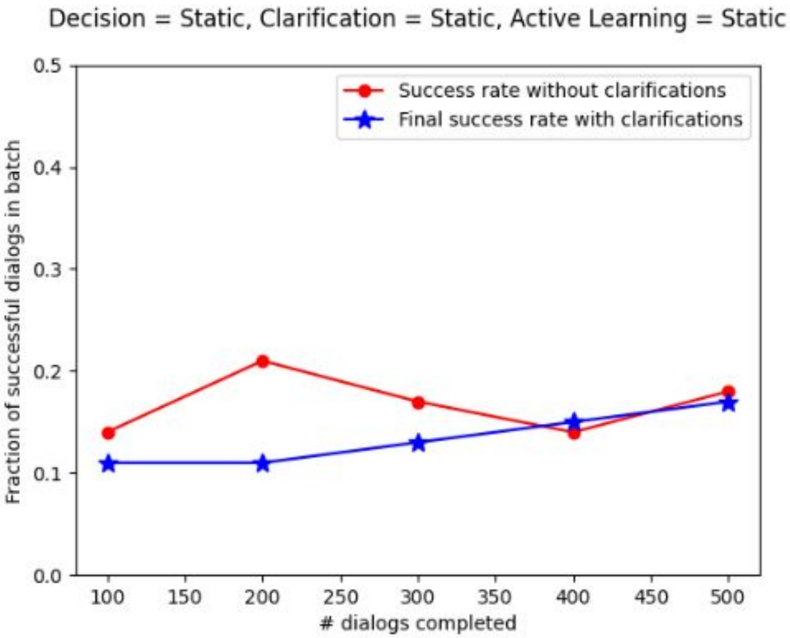


Utility of Clarifications



Utility of Clarifications





Summary

- We train a hierarchical dialog policy to trade off opportunistic active learning, attribute based clarification and task completion in a language based image retrieval task.
- Our learned policy is more successful than a static baseline while using fewer dialog turns on average.
- In our task setup, both good clarifications and active learning queries are necessary to improve performance over direct retrieval.

Outline

- Dialog Policy Learning for Joint Clarification and Active Learning Queries
  - Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
  - Human Evaluation
  - Extension to Joint Embedding Based Grounding Model

Human Evaluation - Experiment Changes

- Descriptions from human users contained far fewer attributes than product titles
- Changes in task setup -
  - Provide one attribute from product title as simulated description
  - Smaller and easier active test set

Experiment Interface

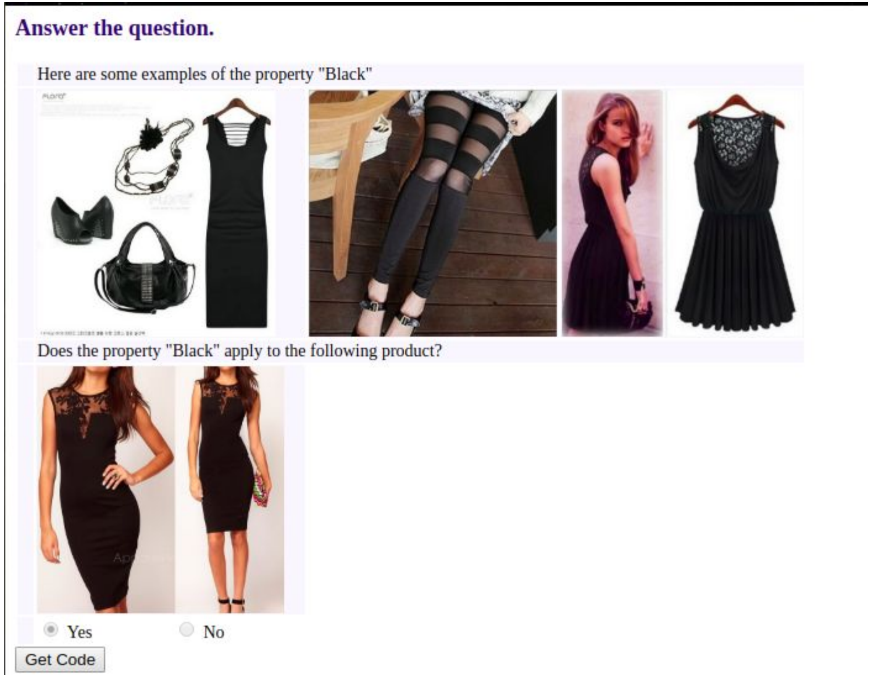
Describe the product in the image.

Describe the product in the image.

red dress

Continue

Experiment Interface



Experiment

- Initialization, training and test phases run in new simulated setup
- Run a single batch of interactions on Amazon Mechanical Turk with final policy and classifiers

Results

Policy	Simulation – Fraction of Successful Dialogs	Simulation – Average Dialog Length
Static	0.23	20.0
Learned	**0.65**	20.0

The learned policy is considerably more successful in the new simulated setup but is unable to shorten dialogs compared to the baseline.

Results

Policy	Simulation – Fraction of Successful Dialogs	Simulation – Average Dialog Length	AMT – Fraction of Successful Dialogs	AMT – Average Dialog Length
Static	0.23	20.0	0.06	19.16
Learned	**0.65**	20.0	*0.16*	18.86

- The performance of both policies drops in AMT interactions.
- The learned policy is still somewhat more successful (p <= 0.1)

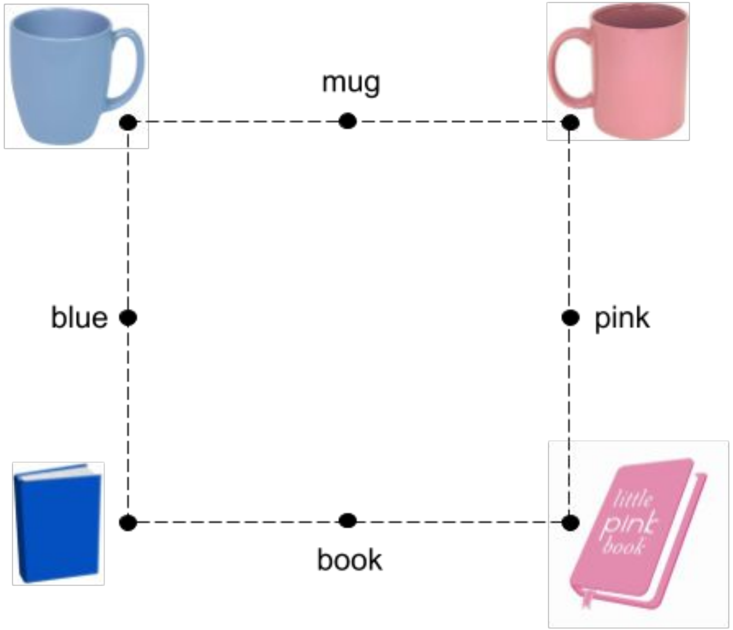
Outline

- Dialog Policy Learning for Joint Clarification and Active Learning Queries
  - Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
  - Human Evaluation
  - Extension to Grounding Model Based on Joint Embeddings

Motivation

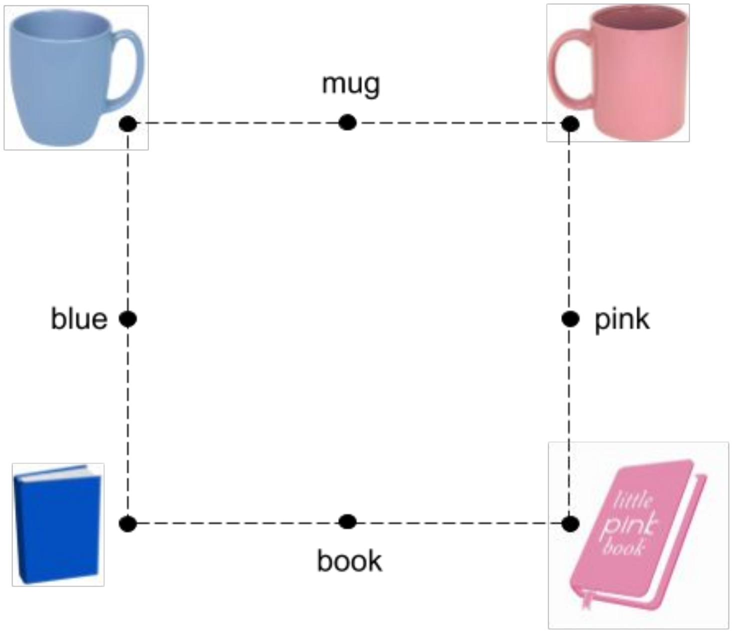
- Independent classifiers cannot identify correlations between properties
- Multilabel classifiers assume a fixed set of properties

Grounding Model



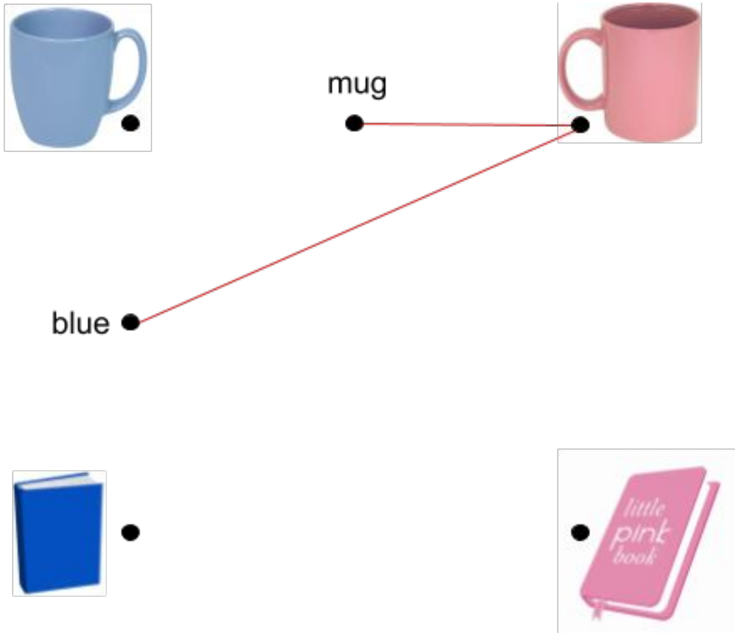
Grounding Model

- Represent words and images as vectors in the same space.
- Words are near images they apply to and vice versa.



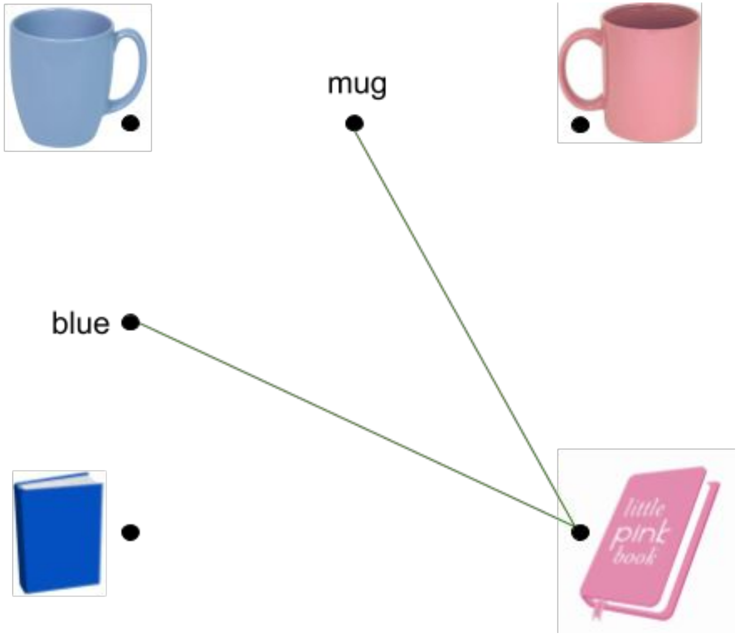
Grounding Model

To ground a description, such as “blue mug”, find the image which minimizes the sum of distances to the words.



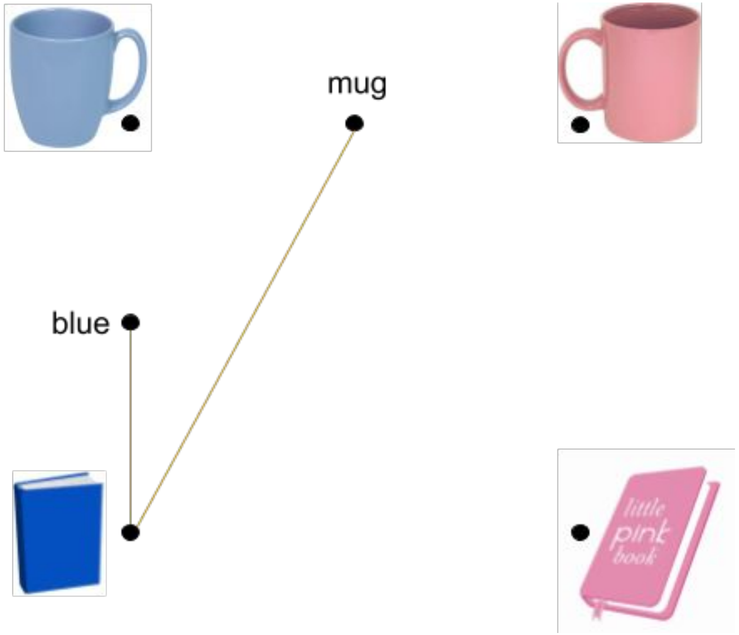
Grounding Model

To ground a description, such as “blue mug”, find the image which minimizes the sum of distances to the words.



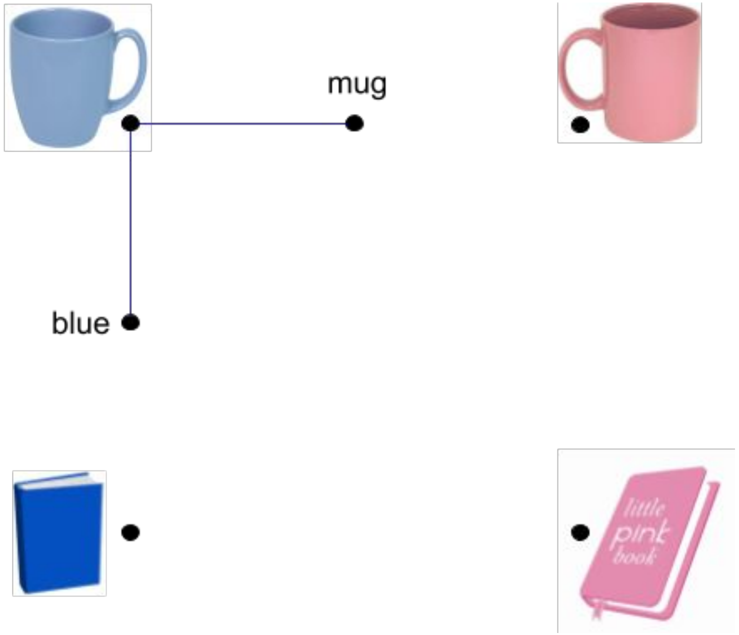
Grounding Model

To ground a description, such as “blue mug”, find the image which minimizes the sum of distances to the words.

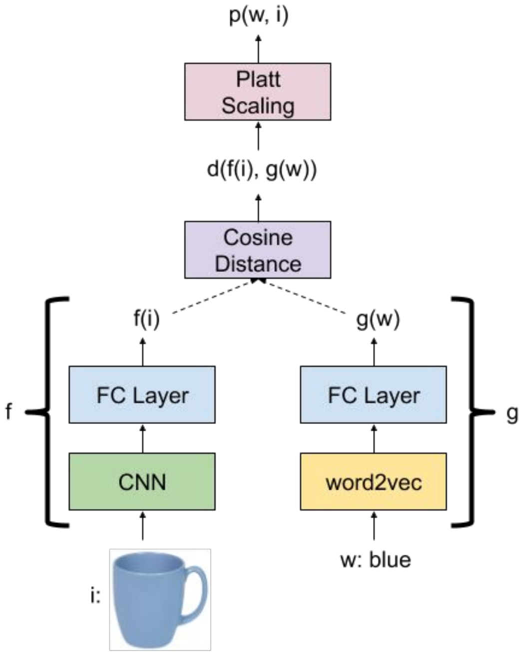


Grounding Model

To ground a description, such as “blue mug”, find the image which minimizes the sum of distances to the words.



Grounding Model



Grounding Model

d(f(), g(*blue*)) $\leq$ d(f(), g(*blue*))

d(f(), g(*blue*)) $\leq$ d(f(), g(*pink*))

- Constraints captured using a ranking loss
- Platt scaling parameters are trained using log loss

Preliminary Results

Clarifications with a high estimate of information gain do not necessarily increase the belief of the correct target image.

Discussion

Possible reasons why our estimate of information gain is not able to identify helpful clarifications -

- Noise in annotations used to provide responses
- Grounding model does not produce a true probability distribution

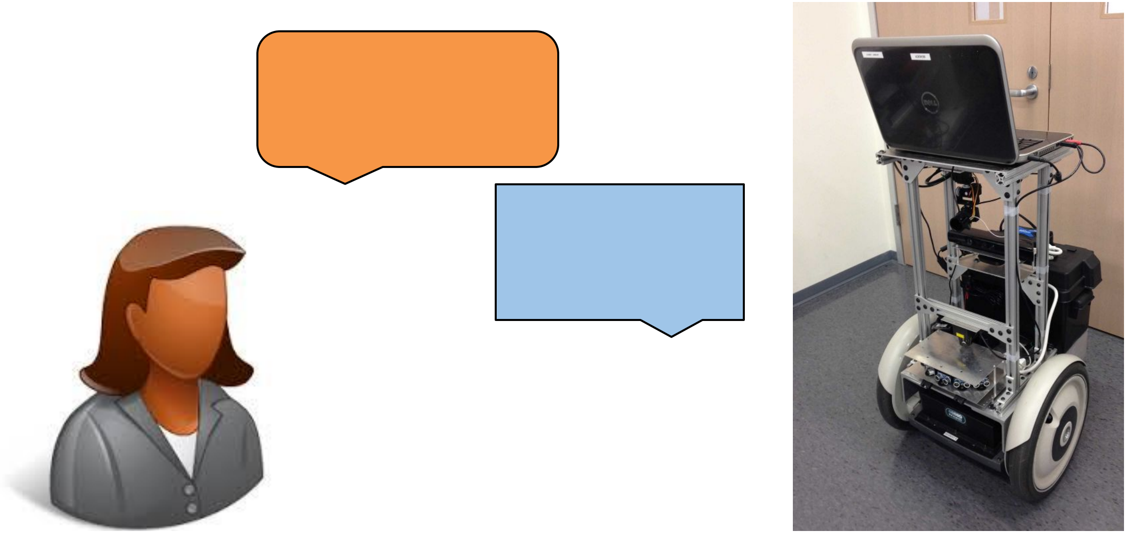
Future Work

- Better learned spaces - Possibly using pretrained models such as ViLBERT, LXMERT
- Techniques such as adversarial loss to make the learned space smoother.

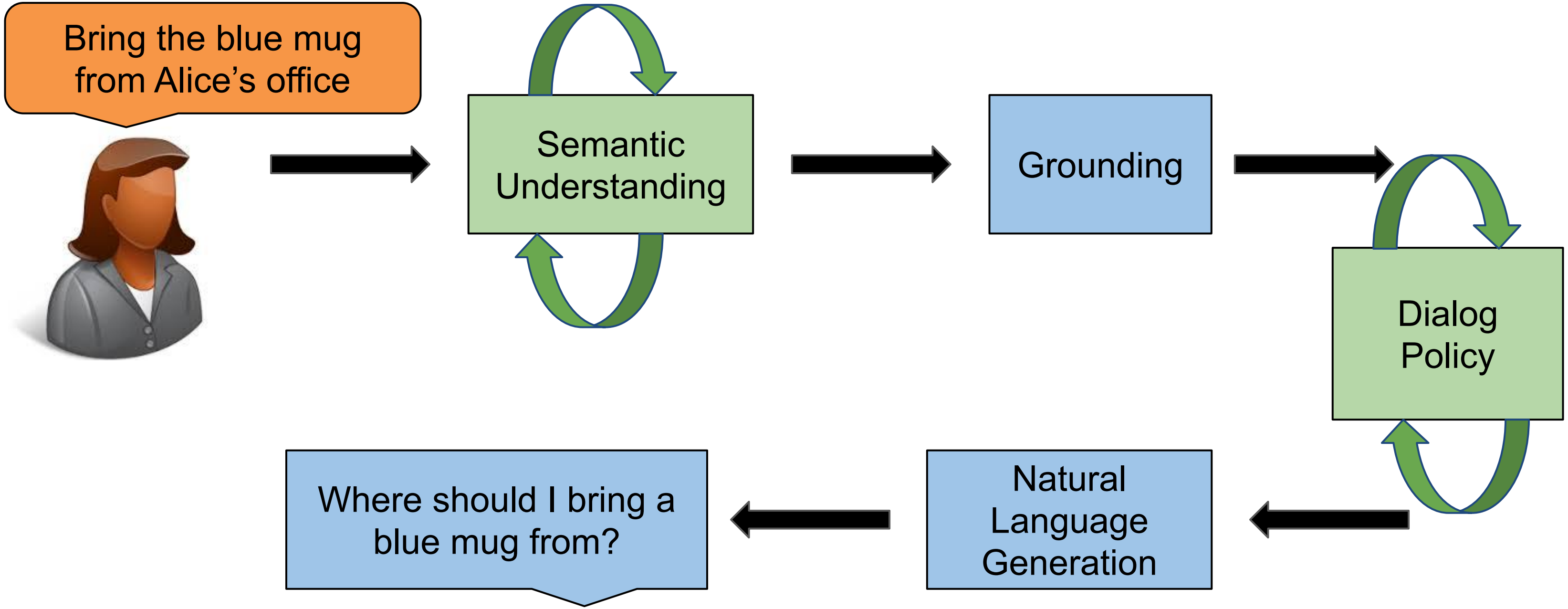
Outline

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

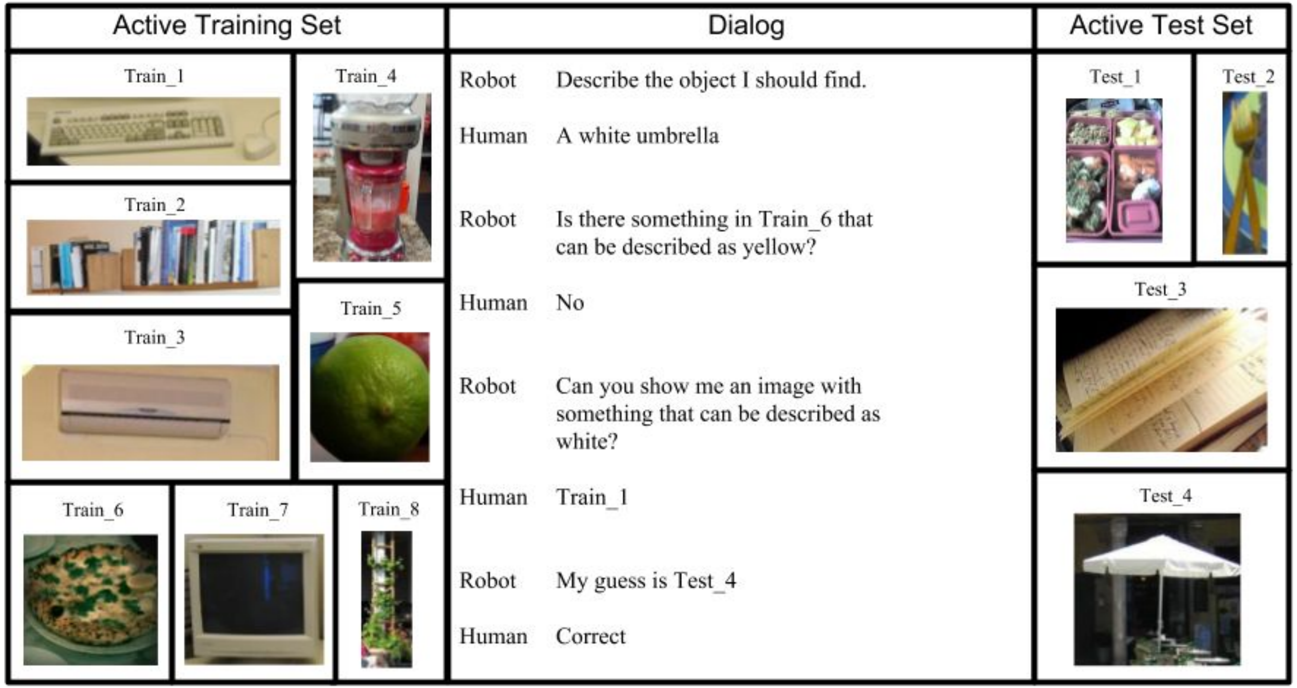
Dialog as a Vehicle for Lifelong Learning of Grounded Language Understanding Systems



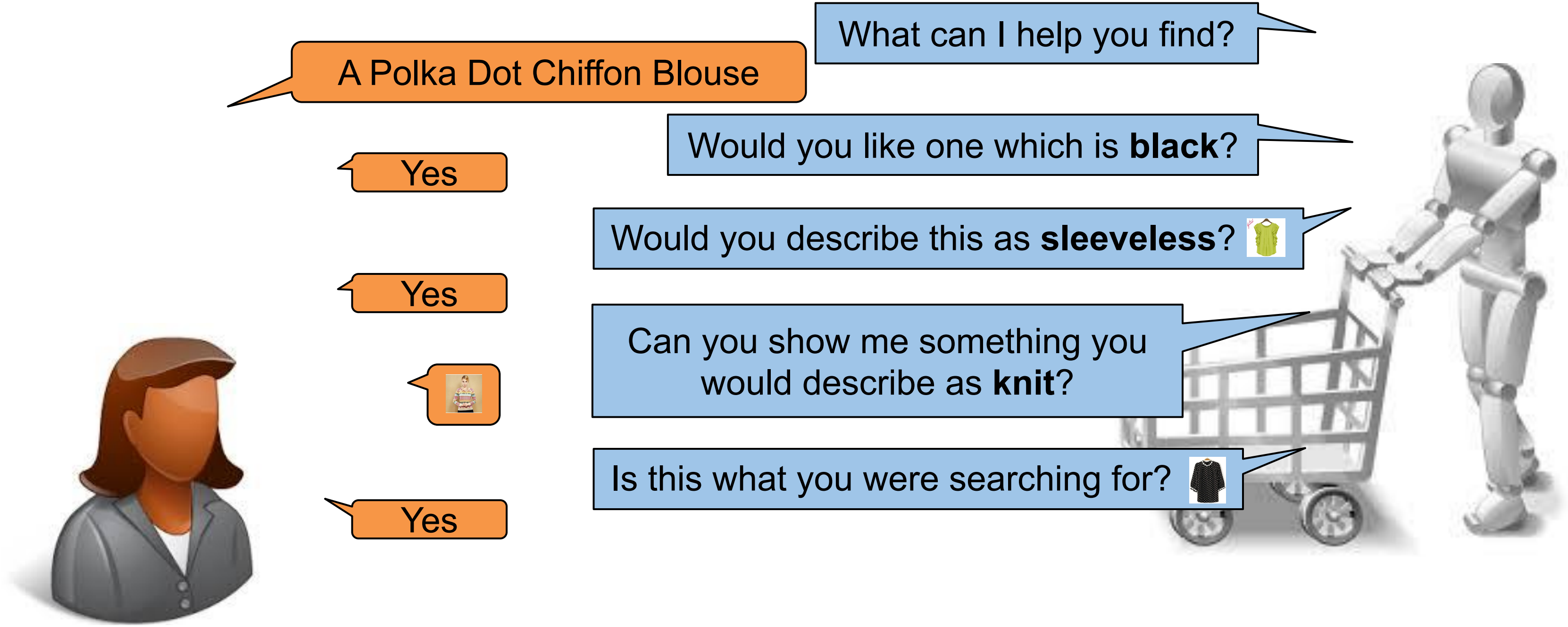
Joint Parser and Policy Learning



Policy Learning for Opportunistic Active Learning



Dialog Policy Learning for Joint Clarification and Active Learning Queries



Outline

- Background
- Integrating Learning of Dialog Strategies and Semantic Parsing (Padmakumar et.al., 2017)
- Opportunistic Active Learning for Grounding Natural Language Descriptions (Thomason et. al., 2017)
- Learning a Policy for Opportunistic Active Learning (Padmakumar et. al., 2018)
- Dialog Policy Learning for Joint Clarification and Active Learning Queries (Padmakumar and Mooney, in submission)
- Summary
- New Directions (Padmakumar and Mooney, RoboDial 2020)

Dialog as a Vehicle for Lifelong Learning

[Padmakumar and Mooney, RoboDial 2020]

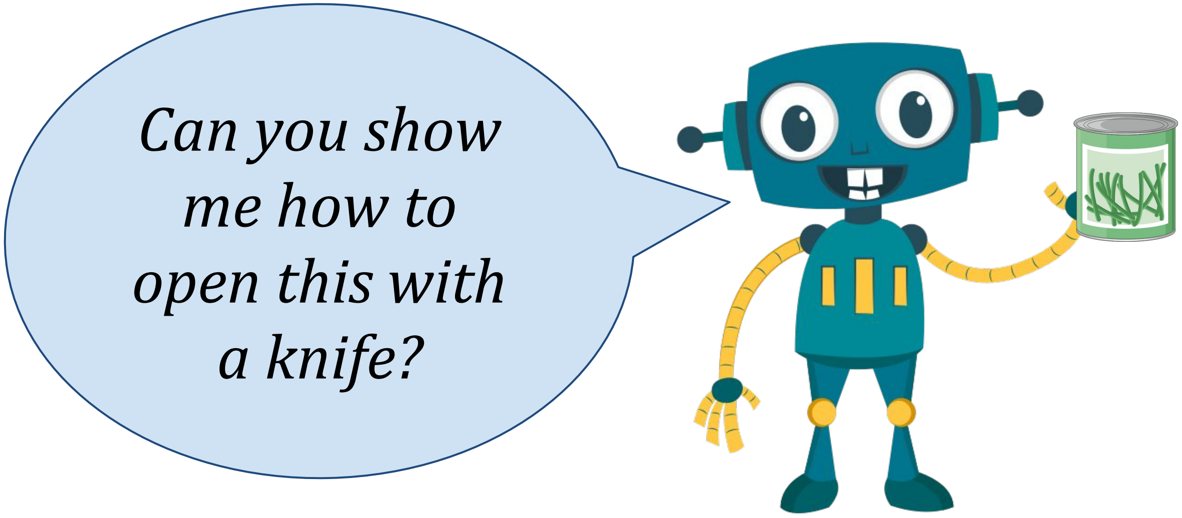
- New challenge area for dialog researchers
- Goal: Design dialog systems that can better support lifelong learning

Challenges: Active Learning

- Improving sample complexity
- Few shot adaptation of pretrained models
- Better robustness and transferability of RL policies for active learning

Challenge: Dialog Act Design

Design new dialog acts that collect labeled data or combine this with task-completion objectives



Challenges: Dataset Collection and Simulation

- Designing simulations to answer a wide range of queries.
- Providing “correct” answers in simulation.
- Sim2Real Transfer

Challenges: User Experience

- Prosodic analysis to identify urgency, stress, sarcasm and frustration in users to determine when it is appropriate to include or avoid data collection queries.
- Demonstrating few-shot learning to keep users motivated.

Dialog as a Vehicle for Lifelong Learning of Grounded Language Understanding Systems

Aishwarya Padmakumar

Doctoral Dissertation Defense