COS 484: Natural Language Processing

L19: Language Grounding - 2

Spring 2022
Logistics

- Sign up for project meetings on April 19
  - Mandatory for every team to meet with your staff guide
- Fill up preference form for poster session on April 21
Some grounding tasks

- **Vision**
  - Captioning
  - Visual question answering (VQA)
  - Spatial reasoning

- **Interaction**
  - Instruction following
  - Text-based games
  - RL for NLP
Instruction Following

- Want to be able to follow instructions in a virtual environment
- “Go along the blue hall, then turn left away from the fish painting and walk to the end of the hallway”

Grounding language to actions

(MacMahon et al., 2006)
Instruction Following

- Train semantic parser on (utterance, action) pairs

**Instruction:** “Go away from the lamp to the intersection of the red brick and wood”

**Basic:**
- Turn ( ),
- Travel ( steps: 1 )

**Landmarks:**
- Turn ( ),
- Verify ( left: WALL, back: LAMP, back: HATRACK, front: BRICK HALL),
- Travel ( steps: 1 ),
- Verify ( side: WOOD HALL )

*(Chen and Mooney, 2011)*
Grounding semantics in control applications

1. Use feedback from task to understand language

2. Use language to improve performance in control applications

Alleviate dependence on supervised annotation

Score: 7
Score: 107
Reinforcement learning
Reinforcement Learning

• Delayed feedback

  ⇒ How to perform credit assignment for individual actions

• Large number of possible action sequences
  ⇒ Need for effective exploration

Improved language understanding translates to improved task performance
Playing Civilization by reading game manuals

Neural network for policy

Settlers unit, candidate action 1: \textit{irrigate}

Features:
- action = \textit{irrigate} and action-word = "irrigate"
- action = \textit{irrigate} and state-word = "land"
- action = \textit{irrigate} and terrain = plains
- action = \textit{irrigate} and unit-type = settler
- state-word = "city" and near-city = true

Settlers unit, candidate action 2: \textit{build-city}

Features:
- action = \textit{build-city} and action-word = "irrigate"
- action = \textit{build-city} and state-word = "land"
- action = \textit{build-city} and terrain = plains
- action = \textit{build-city} and unit-type = settler
- state-word = "city" and near-city = true

Relevant text: "Use settlers to irrigate land near your city"

Predicted action words: "irrigate", "settler"

Predicted state words: "land", "near", "city"

<table>
<thead>
<tr>
<th>Method</th>
<th>% Win</th>
<th>% Loss</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>0</td>
<td>100</td>
<td>—</td>
</tr>
<tr>
<td>Built-in AI</td>
<td>0</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>Game only</td>
<td>17.3</td>
<td>5.3</td>
<td>± 2.7</td>
</tr>
<tr>
<td>Latent variable</td>
<td>26.1</td>
<td>3.7</td>
<td>± 3.1</td>
</tr>
<tr>
<td><strong>Full model</strong></td>
<td><strong>53.7</strong></td>
<td><strong>5.9</strong></td>
<td>± 3.5</td>
</tr>
<tr>
<td>Randomized text</td>
<td>40.3</td>
<td>4.3</td>
<td>± 3.4</td>
</tr>
</tbody>
</table>

(Branavan et al., 2012)
Learning a grounding

- How do we map symbols in language (i.e. words) to entities and concepts in the world?
- Can an agent learn grounding through interaction

[Grounding Language to Entities and Dynamics for Generalization in Reinforcement Learning. Austin W. Hanjie, Victor Zhong, Karthik Narasimhan; ICML 2021]
• Agent can move around and interact in a simulated environment

• Receives global state observations, rewards

• Has access to a text "manual" describing entities and dynamics, throughout an episode

• Agent is not provided any prior mapping between the observations (alion) and symbols in text (wizard, mage) to help it "read" the manual.

[Grounding Language to Entities and Dynamics for Generalization in Reinforcement Learning. Austin W. Hanjie, Victor Zhong, Karthik Narasimhan; ICML 2021]
Messenger

- Multi-game benchmark with separate train and test splits
- In each game, agent has to first pick up a *message*, and deliver it to *goal* entity, while avoiding an *enemy*
- Each game has different entities, each with different roles and different dynamics
  - There may be multiple entities of the same type! (e.g. mage in game 1)
- The agent must consult a natural language manual in order to consistently win
  - Manual may contain extraneous/incorrect information (e.g. point 6 here).
Messenger: Statistics

• Random instantiation of roles each time
• 44/32/32 train/val/test game variants
• 5000+ textual descriptions, vocabulary size of 1125
• 30-60 words/manual, completely human written (crowdsourced)

GAME 1 MANUAL

1. at a particular locale, there exists a motionless mongrel that is a formidable adversary.
2. the top-secret paperwork is in the crook’s possession, and he’s heading closer and closer to where you are.
3. the crucial target is held by the wizard and the wizard is fleeing from you.
4. the mugger rushing away is the opposition posing a serious threat.
5. the thing that is not able to move is the mage who possesses the enemy that is deadly.
6. the vital goal is found with the canine, but it is running away from you.
Why is Messenger challenging?

- Agent has to learn an accurate grounding purely through interaction
- Wide variation in how an entity is described - e.g. use of multiple synonyms ('crook, thief'), non-templated freeform text
- No overlap in terms of entity-role-dynamics combinations between train and test games

“The top-secret paperwork is in the crook’s possession, and he’s heading closer and closer to where you are"
Our model: Entity Mapper with Multimodal Attention (EMMA)

Jointly process observations with text manual for control policy
EMMA does better on Messenger...

Win rates on stage 2 of Messenger for baselines

- Mean BOS
  - Train: 2.10
  - Test: 4.70
- Bayesian Attention
  - Train: 69.00
  - Test: 41.00
- txt2pi (Zhong et al)
  - Train: 94.00
  - Test: 0.30
- EMMA
  - Train: 95.00
  - Test: 85.00
... but some stages continue to prove challenging

Win rates on stage 3 of Messenger for baselines

<table>
<thead>
<tr>
<th>Method</th>
<th>Train</th>
<th>Test</th>
<th>Train</th>
<th>Test</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean BOS</td>
<td>0.00</td>
<td>0.00</td>
<td>1.40</td>
<td>2.70</td>
<td>3.00</td>
<td>2.60</td>
</tr>
<tr>
<td>Bayesian Attention</td>
<td>22.00</td>
<td>10.00</td>
<td>22.00</td>
<td>10.00</td>
<td>22.00</td>
<td>10.00</td>
</tr>
<tr>
<td>txt2pi (Zhong et al)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMMA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EMMA learns to map each description to the entity it describes.
You are standing in an open field west of a white house, with a boarded front door. There is a small mailbox here.

- open mailbox
- go east
- search field

Underlying game state (h1)

(Narasimhan et al., 2015)
You are standing in an open field west of a white house, with a boarded front door. There is a small mailbox here.

*open mailbox*
Opening the mailbox reveals a leaflet.

Underlying game state (h2)
You are standing in an open field west of a white house, with a boarded front door. There is a small mailbox here.

You are in an open field next to a white house. The house’s front door is boarded shut. You see a small mailbox here.
Opportunity

Grounded language learning

You are standing in an open field west of a white house, with a boarded front door. There is a small mailbox here.

♦ open mailbox
Opportunity

Grounded language learning

In-game rewards provide unstructured feedback

+10 gold  +5 health
Opportunity

Learn language while performing tasks

In-game rewards provide unstructured feedback to learn

+10 gold

+5 health
You are standing in …

Recurrent Neural Network

You  are  standing  in  …

(vector representation)

(hidden layer)

(input text)

(Narasimhan et al., 2015)
LSTM-DQN: Action Scorer

Deep Neural Network for control policy

Input text \( T \) → Recurrent Neural Network \( u \) \( \rightarrow \) Output text \( Q \) \( \rightarrow \) \( Q(s, a) \)

Learn parameters using Q-learning

Recurrent Neural Network to map text to vector representation
Results

<table>
<thead>
<tr>
<th>Quest completion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM-DQN</td>
</tr>
<tr>
<td>BI-DQN</td>
</tr>
<tr>
<td>BOW-DQN</td>
</tr>
<tr>
<td>Random</td>
</tr>
</tbody>
</table>

- LSTM-DQN: 98.0%
- BI-DQN: 55.0%
- BOW-DQN: 40.0%
- Random: 15.0%
Visualizing Learnt Representations

t-SNE visualization of vectors learnt by agent
Contextual Action Language Model (CALM)

• **Want:** Generate sensible action commands

• **Idea:** Train a *single* language model to generate action candidates for any game

• Actions are subsequently reranked by an RL agent using game-specific rewards

*Observation:* You are in the living room. There is a doorway to the east, a wooden door with strange gothic lettering to the west, which appears to be nailed shut, a trophy case, and a large oriental rug in the center of the room. You are carrying: A brass lantern . . .

*Random Actions:*
- close door, north a, eat troll with egg, …

*CALM (n-gram) Actions:*
- enter room, leave room, lock room, open door, close door, knock on door, …

*CALM (GPT-2) Actions:*
- east, open case, get rug, turn on lantern, move rug, unlock case with key, …

*Next Observation:* With a great effort, the rug is moved to one side of the room, revealing the dusty cover of a closed trap door...

(Yao et al., 2020)
Semantics does not exist in isolation

Diagram:
- Language
  - Logical forms
  - Parse trees
  - Vector representations
- Perception
- Interaction
Is coffee a carcinogen?

Coffee significantly reduced ER and cyclin D1 abundance in ER(+) cells

... Coffee reduced the pAkt levels in both ER(+) and ER(-) cells.
Information Extraction: State of the Art

Dependence on large training sets

ACE: 300K words
Freebase: 24M relations

Not available for many domains (ex. medicine, crime)

Even large corpora do not guarantee high performance
~ 75% F1 on relation extraction (ACE)
~ 58% F1 on event extraction (ACE)
A 2 year old girl and four other people were wounded in a shooting in West Englewood Thursday night, police said.
A 2 year old girl and four other people were wounded in a shooting in West Englewood Thursday night, police said.

The last shooting left five people wounded.
Incorporate external evidence

Traditional formulation

[ Narasimhan et al. 2016 ]

extract + reason

extract

aggregate

find extra articles

extract
Challenges

1. Event Coreference

4 adults, 1 teenager shot in west Baltimore

<table>
<thead>
<tr>
<th>All</th>
<th>News</th>
<th>Shopping</th>
<th>Images</th>
<th>Videos</th>
<th>More</th>
<th>Search tools</th>
</tr>
</thead>
</table>

About 16,200,000 results (0.63 seconds)

4 adults, 1 teenager shot in west Baltimore | Maryland News ... www.wbaltv.com/news/...shot-in-west-baltimore/32156116 | WBAL-TV | Apr 3, 2015 - Five people were shot Thursday afternoon in west Baltimore.

1 killed, 3 injured in Baltimore shooting, police say ... - WBAL www.wbaltv.com/news/...shot-in-west-baltimore/3658266 | WBAL-TV | Nov 21, 2015 - 2 teens, 2 adults shot on Stricker Street ... man was killed and three others were injured in a shooting Saturday morning in west Baltimore, police said. ... Mom tries to buy baby for her 14-year-old daughter; WBALTV.com. Undo.

10-year-old boy shot in West Baltimore - Baltimore Sun www.baltimoresun.com/.../baltimore.../bs-md-ci-shoot... | The Baltimore Sun | Sep 3, 2015 - A 10-year-old boy was shot Thursday night, along with two adult ... Baltimore police report 6 shootings, including one of a teenage boy. ... The homicide occurred about 4:30 p.m. at Ninth and East Jeffrey streets in Brooklyn, police said. ... At 1:20 a.m., officers found a 32-year-old Baltimore man shot in the ... Several irrelevant articles!

2. Reconciling Predictions

**Shooter:** Scott Westerhuis  
**NumKilled:** 4  
**Location:** S.D

**Shooter:** Scott Westerhuis  
**NumKilled:** 6  
**Location:** Platte

Inconsistent extractions
Learning through reinforcement

Start with traditional extraction system

S.D. dad killed wife, four kids with shotgun setting house ablaze and killing self: author

**Shooter:** Scott Westerhuis

**NumKilled:** 4

**Location:** S.D
Learning through reinforcement

Perform a query and extract from a new article

**Shooter**: Scott Westerhuis  
**NumKilled**: 4  
**Location**: S.D

**Shooter**: Scott Westerhuis  
**NumKilled**: 6  
**Location**: Platte
Learning through reinforcement

Current

State

New

- Shooter: Scott Westerhuis
  - NumKilled: 4
  - Location: S.D

- Shooter: Scott Westerhuis
  - NumKilled: 6
  - Location: Platte
1. **Reconcile (d)** old values and new values.
   - Pick a single value, all values or no value from new set
2. Decide how to proceed:
   - Stop
2. Decide how to proceed:
+ Select next query (q)
Acquiring external evidence

1. Select a query to search for articles on the same event

   ![Search Query](shooting_in_platte_september_2015)

2. Use base extractor to obtain values for entities of interest

   - **Shooter**: Scott Westerhuis
   - **NumKilled**: 6
   - **Location**: Platte

3. Reconcile old and new extractions

   - **Old Extractions**
     - **Shooter**: Scott Westerhuis
     - **NumKilled**: 4
     - **Location**: S.D
   - **New Extractions**
     - **Shooter**: Scott Westerhuis
     - **NumKilled**: 6
     - **Location**: Platte
Learning from rewards

• Change in accuracy

\[
R(s, a) = \sum_{\text{entity } j} \text{Acc}(e_j^{\text{cur}}) - \text{Acc}(e_j^{\text{prev}}) = 1
\]

• Small penalty for each transition

Previous Values

Shooter: Scott Westerhuis
NumKilled: 6
NumWounded: 1
Location: Platte

Current Values

Shooter: Scott Westerhuis
NumKilled: 6
NumWounded: 0
Location: Platte

Shooter:
NumKilled:
NumWounded:
Location:
Mass shootings in the United States

Adulteration incidents from Foodshield EMA

~300 training instances
Accuracy

NumKilled

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base extractor</td>
<td>69.7</td>
</tr>
<tr>
<td>Confidence Agg.</td>
<td>70.3</td>
</tr>
<tr>
<td>Meta-Classifier</td>
<td>70.7</td>
</tr>
</tbody>
</table>
Sequential decision making helps!