Learning to Follow Navigational Directions
Adam Vogel and Dan Jurafsky
Stanford University

State s
Current utterance: "then eh go right until you’re between springbok boks and highest viewpoint"
Current Path: (springboks, south)

Action a
Target Landmark & Side: (springboks, south)

Reward R(s,a)
Order: correct
Side: wrong
Order: correct
Side: correct

Reinforcement Learning
Find θ which best models expected reward of choosing an action:

Q(s,a) = θ•φ(s,a)

Learning Algorithm: SARSA, an approximation to value iteration

Map Task
• Two people have a map with named landmarks
• The instruction giver describes a path on the map
• The instruction follower draws the path described
• Our system learns to be the instruction follower

Directions
1. right go vertically down until you’re underneath eh diamond mine
2. then eh go right until you’re between springbok boks and highest viewpoint
3. right go right ah right to your right
4. then up around the uh highest viewpoint
5. eh once you’re once you’ve eh passed it right go vertically down to just above eh safari truck

Evaluation
Baseline policy: visit the closest unvisited landmark next

<table>
<thead>
<tr>
<th></th>
<th>Order F</th>
<th>Side F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>32.2</td>
<td>52.2</td>
</tr>
<tr>
<td>SARSA</td>
<td>48.2</td>
<td>61.2</td>
</tr>
</tbody>
</table>

Features φ(s,a)
• Is the target landmark is closest?
• Is this the closest side of the landmark?
• Is this the null action?
• How many words in the name of the target landmark occur in the utterance?
• Allocentric spatial features: A feature combining each spatial word in the utterance with the side we pass the landmark on. These features capture landmark-centric spatial relations.
  Ex: right appears and we pass to the south of the landmark
• Egocentric spatial features: A feature for every spatial word in the utterance conjoined with the cardinal direction we move in. These capture agent-centric spatial relations.
  Ex: right appears and we move to the east

Acquired Spatial Semantics
Above, Below, Right, South, Top

This figure shows the normalized feature weights for the allocentric and egocentric spatial relations. For example, the top left shows that if above occurs in an utterance, our model prefers to pass the landmark to the north.

Learning by Apprenticeship
Problem: Without linguistic annotation:
• Can we learn the meaning of spatial words like *left* and *down*?
• Can we learn to interpret textual directions on a map?

Traditional approaches to semantic interpretation are insufficient:
• Supervised approaches use expensive hand-labeled semantic annotation.
• Unsupervised approaches that learn semantics only from word distributions can’t deal with spatial terms.

Our solution: Induce the meaning of spatial language from natural language route descriptions paired with a path on a map, using reinforcement learning.