Panning for Gold: Finding Relevant Semantic Content for Grounded Language Learning
David L. Chen and Raymond J. Mooney
The University of Texas at Austin

Sample Instructions
• Take your first left. Go all the way down until you hit a dead end.
• Go towards the coat hanger and turn left at it. Go straight down the hallway and the dead end is position 4.
• Walk to the hat rack. Turn left. The carpet should have green octagons. Go to the end of this alley. This is p-4.
• Walk forward once. Turn left. Walk forward twice.

Observed Action Sequence
Forward, Left, Forward, Forward

Formal Definition
Given:
\{ (e_1, a_1, w_1), (e_2, a_2, w_2), \ldots, (e_n, a_n, w_n) \}
e_i – A natural language instruction
a_i – An observed action sequence
w_i – A world state
Goal:
Find the correct plan p_i corresponding to the instruction e_i

Plan Construction
Basic plan: Directly model the observed actions
Landmarks plan: Add interleaving verification steps

Plan Refinement
• First learn a lexicon. To learn the meaning of the word/short phrase w:
  1. Collect all plans that co-occur with w and add them to MeaningSet(w)
  2. Repeatedly take intersections of all possible pairs of members of MeaningSet(w) and add any new entries to MeaningSet(w)
  3. Rank the entries by the scoring function: \( Score(w, g) = p(g|w) - p(g|\neg w) \)

Use the learned lexicon to help remove extraneous components of the graph
Instruction: Turn left and walk to the sofa

Refined landmarks plan:

Data
Adapted from data collected by MacMahon et al. (2006)
# instructions: 3236
Vocabulary size: 629
Avg. # words: 7.8 (Std. Dev. 5.1)
Avg. # actions: 2.1 (Std. Dev. 2.4)

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic plans</td>
<td>81.47</td>
<td>56.04</td>
<td>66.40</td>
</tr>
<tr>
<td>Landmarks plans</td>
<td>45.39</td>
<td>85.56</td>
<td>59.31</td>
</tr>
<tr>
<td>Refined landmarks plans</td>
<td>80.59</td>
<td>77.49</td>
<td>79.01</td>
</tr>
<tr>
<td>Refined landmarks plans (no temporal links)</td>
<td>80.54</td>
<td>68.87</td>
<td>74.25</td>
</tr>
</tbody>
</table>

Partial matching accuracy compared to human annotated plans