Alignment-Based Compositional Semantics for Instruction Following

Jacob Andreas and Dan Klein
UC Berkeley
Following instructions

*go down the green hall*
*go all the way*
*then face the chair*
Following instructions

go down the green hall

go all the way

then face the chair
Following instructions: parsing

go down the **green hall**
go all the way
then face the chair
Following instructions: parsing

go down the green hall

go all the way

then face the chair
go down the green hall

go all the way

then face the chair
Following instructions: planning

*go down the green hall*

*go all the way*

*then face the chair*
Following instructions

Planning
(no instruction)

[ go down the hall, go all the way ]

Parsing

green hall

face the chair
Previous work

Planning
- Branavan et al. (2009)
- Vogel and Jurafsky (2010)

Parsing
- Chen and Mooney (2011)
- Artzi and Zettlemoyer (2013)
- Kim and Mooney (2013)
Parsing and planning

**Sequence-to-sequence**

*go down the green hall*  *turn left*

**Structure-to-structure**

*go down*  *the*  *green*  *hall*
SEQUENCE-TO-SEQUENCE MODEL
go down the yellow hall  

turn left
go down the yellow hall  

[Alignments]

turn left
go down the yellow hall  turn left

Sequence-to-sequence

Alignments

Actions
Sequence-to-sequence

Go down the yellow hall  Turn left

Alignments

Actions
go down the yellow hall turn left

Sequence-to-sequence

Alignments

Actions
go down the yellow hall  turn left

Sequence-to-sequence

Alignments

Actions
go down the yellow hall  turn left
go down the yellow hall  turn left
go down the yellow hall   turn left
Sequence-to-sequence

go down the yellow hall  turn left
Sequence-to-sequence

go down the yellow hall  turn left
Sequence-to-sequence

go down the yellow hall  turn left
STRUCTURE-TO-STRUCTURE MODEL
go down the yellow hall
Structure-to-structure

go → down → hall → yellow

the
Structure-to-structure

go → down → hall → yellow → the
Structure-to-structure

go down

the yellow hall
Structure-to-structure
go down the yellow hall
Structure-to-structure

the yellow
down
gohall
JOINT MODEL
go down the yellow hall turn left
Joint model

go down the yellow hall  turn left
Joint model

go down the yellow hall

turn left
Joint model

go down the yellow hall

turn left
Joint model

\[ \psi \left( \begin{array}{c} \text{go down} \\ \text{the yellow hall} \end{array} \right) = \theta^\top \left( \phi(go, \bullet) + \phi(down, \bullet) + \cdots \right) \]
Joint model

\[ p(\text{actions, alignment} \mid \text{text}; \theta) \propto \exp \sum \psi \left( \begin{array}{c}
\text{go down} \\
\text{the yellow hall}
\end{array} \right) \]
go down the yellow hall  turn left
LEARNING / INFEERENCE
Learning: coordinate ascent

\[
\max_{\theta, \text{alignments}} p(\text{actions, alignments} \mid \text{text}; \theta)
\]

\[
\max_{\text{alignments}} \quad \text{using custom alignment dynamic program}
\]

\[
\max_{\theta} \quad \text{using L-BFGS}
\]
Decoding: coordinate ascent

\[ p(\text{actions, alignments} \mid \text{text}; \theta) \]

\[ \text{max} \]
\[ \text{alignments} \]
using custom alignment dynamic program

\[ \text{max} \]
\[ \text{alignments} \]

\[ \text{max} \]
\[ \text{actions} \]
using a planner
EXPERIMENTS
Three tasks
Puzzle solving

clear the two long columns, and then the row

[Branavan+ 09]
Puzzle solving

clear the two long columns, and then the row
Puzzle solving

Task completion (%)

- No text: 78%
- This work: 86%
you should see a grey floor to your right and a chair in front of you
<table>
<thead>
<tr>
<th>Task completion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
</tr>
<tr>
<td>55</td>
</tr>
<tr>
<td>45</td>
</tr>
<tr>
<td>35</td>
</tr>
</tbody>
</table>
right round the white water [...] but stay quite close ’cause you don’t otherwise you’re going to be in that stone creek

[Anderson+ 91]
Map reading

Plan execution ($F_0$)

0
15
30
45
60
Ablations: maze navigation

Task completion (%)

60
This work
Ablations: map reading

Plan execution ($F_0$)

Without grounding graphs

This work
Conclusion

Structured alignment/decoding gives us best aspects of:

- Compositional semantics (like a parser)
- Sequence structure (like a planner)
THANK YOU