Grounding Antonym Adjective Pairs through Interaction

Maxwell Forbes
Michael Jae-Yoon Chung
Maya Cakmak
Luke Zettlemoyer
Rajesh P. N. Rao
facial expression and gesture recognition

natural language generation

symbol grounding

semantic language analysis

speech recognition
Symbol Grounding

“mailbox”
“mailbox”
“mailbox”
[ADJECTIVES]

“heavy” compared to what?
“heavy”
"heavy"

sensor_1: \(x\) 
\[5 < x < 10\]

sensor_2: \(f_1, f_2\) 
\[f_1 + f_2 > 4\]
Grounding Antonym Adjective Pairs

“heavy” vs “light”

“bright” vs “dark”

“near” vs “far”

...
## Why should we care?

<table>
<thead>
<tr>
<th>Specificity</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Bring me the <strong>heavier</strong> mug.&quot;</td>
<td>&quot;Please pick up that plate.&quot;</td>
</tr>
<tr>
<td>&quot;What’s the <strong>brightest room</strong> in my house?&quot;</td>
<td>&quot;It is <strong>too far</strong> away.&quot;</td>
</tr>
</tbody>
</table>
Related Work

Grounding of Word Meanings in Multimodal Concepts Using LDA
Nakamura et al. 2009

On the Integration of Grounding Language and Learning Objects
Yu et al. 2004
Our Goals

Interactive

Space Between Adjectives

"heavy" → $S_1 = 0.5$, $f_1 > 4$, ...

"light" → $S_1 > 0.5$, $f_1 = ?$, ...


Why Interactive?
Why Space Between Adjectives?
my sensor features are: 

\[
\begin{align*}
    f_1 &= 0.54627, \\
    f_2 &= 0.75788, \\
    \ldots
\end{align*}
\]

heavy vs light

“heavy”

my sensor features are: 

\[
\begin{align*}
    f_1 &= 0.54627, \\
    f_2 &= 0.75788, \\
    \ldots
\end{align*}
\]

bright vs dark

“dark”
Inside a Classifier

"heavy"

"light"

(some sensor)
More Sensors

"heavy"

"light"

sensor 1

sensor 2

sensor 3
Naïve Bayes

"heavy" vs "light"

features

f1 f2 f3 f4 f5

Gaussian naïve Bayes from scikit-learn
“Relevant” features

“heavy”
vs
“light”

features

f1
f2
f3
f4
f5
Many Classifiers

- x vs y
  - f1
  - f2
  - f3
  - f4
  - f5

- a vs b
  - f1
  - f2
  - f3
  - f4
  - f5

- z vs w
  - f1
  - f2
  - f3
  - f4
  - f5

...
Implementation
Modalities and Features

- **Foot Weights**
  - Left
  - Combined
  - Right

- **Distance** (sonar)
  - Left
  - Right

- **Vision**
  - Brightness
  - # faces
  - Avg mvmt
  - Tot mvmt

- **Haptic** (head sensors)
  - Front
  - Middle
  - Rear
Antonym Adjective Pairs Grounded

1. “heavy” vs “light”
2. “bright” vs “dark”
3. “crowded” vs “lonely”
4. “near” vs “far”
5. “fast” vs “slow”
6. “uncovered” vs “touching”
Between the Adjectives

\[ p(\text{heavy}) = 1.0 \]

\[ p(\text{light}) = 1.0 \]

- "Heavy"
- "More heavy than light"
- "Neither light nor heavy"
- "More light than heavy"
- "Light"
Challenges: NAO Sensors
Limitations

Ambiguous words

“light” vs “bright”

Non-informative phrases

“medium”

“sort of x”
Future Work

Beyond pairs

- very heavy
- heavy
- medium
- light
- feather-light

Systematic evaluation

- user studies: accuracy and robustness

Natural Language Patterns

- meaning of "very"
Thank you