Conceptualizing verbs, nouns and adjectives

Sinan Kalkan
KOVAN Research Lab.
Dept. of Computer Eng.
Middle East Technical University
Ankara, Turkey

www.kovan.ceng.metu.edu.tr
Who am I?

• Leading KOVAN Research Lab with two other faculty members:
  - Erol Şahin
  - Göktürk Üçoluk

Vision
• Border Ownership
• Depth Prediction
• Feature Extraction
• Biometric Identification
• Image Retrieval

Cognitive Robotics
• Conceptualization & Affordances
• Multiple-Levels of Abstraction
• Context

www.kovan.ceng.metu.edu.tr
Introduction: Problem

- give me the tall cup
- What is a cup?
  What is tall?
  What does it mean to give?

Verbs
- give

Adjectives
- tall

Nouns
- cup
Introduction: Concepts

• Rule-Based
Apple = \{\text{color} = \text{“green” AND shape} = \text{“round”}, \ldots\}\n
• Prototype-Based
\[
\text{APPLE} = \left\{
\begin{array}{l}
\text{colour} \quad \text{RED, YELLOW or GREEN respectively} \\
50\%, 25\% \text{ and } 25\% \text{ of the cases} \\
\text{shape} \\
\ldots
\end{array}
\right\
\]

• Exemplar-Based
\[
\text{APPLE} = \left\{ \begin{array}{c}
\text{\includegraphics[width=0.2\textwidth]{apple1}} \\
\text{\includegraphics[width=0.2\textwidth]{apple2}} \\
\text{\includegraphics[width=0.2\textwidth]{apple3}} \\
\ldots
\end{array} \right\}
\]

(Gabora et al., 2008; Kruschke, 2005; Rosch, 1973; Rouder & Ratcliff, 2006)
Introduction:
Affordances and Concepts

Affordance Formalization*

Adding Verb Concepts

Adding Noun/Adjective Concepts

Verb, Noun and Adjective Concepts

Verb Concepts

- Verbs tend to correspond to effect categories rather than single behaviors.

Kalkan et al., accepted.
Dag et al., 2010

Noun/Adjective Concepts

1. Based on visual appearance.
2. Based on what objects afford.

Yuruten et al., 2012; under revision
Atil et al., 2010
Experimental Setup

Kinect: 3D Range Data

Features:
- 3D size, 3D position, surface normal histogram, curvature histogram, presence
Verbs as effect labels
- moved-left (ML)
- moved-right (MR)
- moved-forward (MF)
- pulled (P)
- knocked-down (K)
- disappeared (D)
- no-change (NC)

Nouns and adjectives
- Cup
- Box
- Cylinder
- Ball
- Short-tall
- Thin-thick
- Edgy-round
Experimental Setup:
Behaviors and Effects

**Behaviors**
- Push-left ($PL$)
- Push-right ($PR$)
- Push-forward ($PF$)
- Pull ($PB$)
- Top-grasp ($TG$)
- Side-grasp ($SG$)

**Effects**
- moved-left ($ML$)
- moved-right ($MR$)
- moved-forward ($MF$)
- pulled ($P$)
- knocked-down ($K$)
- disappeared ($D$)
- no-change ($NC$)
Verb Concepts

- grasped
  - ++00--**

- pushed
  - 00*+-*+

Object Concepts

- small
  - +0-**0+

- cup
  - +0-**0+

Behavior Concepts

- push-right

- grasp

Effects

- e
  - e1 e2 e3 e4

- b1 eCK

- ei

- f
  - f1 f2 f3 f4
Conceptualization:
- Capture how features are distributed
Verb Concepts: A prototype-based representation

- NC: No Change
- MR: Moved-right
- ML: Moved-left
- MF: Moved-fwd
- P: Pulled
- K: Knocked
- G: Grasped
- D: disappeared

These prototypes also have mean and standard deviation values of the changes associated with each element.
Comparing conceptualization methods for verbs

1. Prototype-based view 1:
   - Effect prototype-based concepts
   - `+`, `−`, `0`, `*`

2. Prototype-based view 2:
   - Using just the mean & variance of change in features

3. Exemplar-based view:
   - Using all interaction data for comparison

• Comparison using:
  - Recognizing an observed interaction
  - Planning
"what did I do?"

<table>
<thead>
<tr>
<th></th>
<th>No Change</th>
<th>Moved Right</th>
<th>Moved Left</th>
<th>Moved Forward</th>
<th>Pulled</th>
<th>Knocked</th>
<th>Grasped</th>
<th>Disappeared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>390.81</td>
<td>372.24</td>
<td>392.16</td>
<td>389.21</td>
<td>215.56</td>
<td>215.50</td>
<td>392.11</td>
<td>410.31</td>
</tr>
<tr>
<td></td>
<td>392.16</td>
<td>386.92</td>
<td>416.43</td>
<td>416.43</td>
<td>241.06</td>
<td>219.28</td>
<td>395.04</td>
<td>410.31</td>
</tr>
<tr>
<td></td>
<td>237.01</td>
<td>237.42</td>
<td>237.24</td>
<td>237.24</td>
<td>237.42</td>
<td>237.42</td>
<td>237.25</td>
<td>236.84</td>
</tr>
<tr>
<td>Final</td>
<td>731.36</td>
<td>416.42</td>
<td>340.71</td>
<td>393.76</td>
<td>358.06</td>
<td>738.04</td>
<td>790.41</td>
<td>788.84</td>
</tr>
<tr>
<td></td>
<td>732.98</td>
<td>426.71</td>
<td>423.17</td>
<td>428.06</td>
<td>741.11</td>
<td>790.41</td>
<td>788.84</td>
<td>788.84</td>
</tr>
<tr>
<td></td>
<td>789.45</td>
<td>789.83</td>
<td>789.83</td>
<td>789.83</td>
<td>789.83</td>
<td>789.83</td>
<td>789.54</td>
<td>788.84</td>
</tr>
<tr>
<td></td>
<td>925.41</td>
<td>328.75</td>
<td>354.85</td>
<td>928.16</td>
<td>947.51</td>
<td>947.51</td>
<td>947.51</td>
<td>946.21</td>
</tr>
<tr>
<td></td>
<td>929.37</td>
<td>291.77</td>
<td>369.75</td>
<td>929.94</td>
<td>947.01</td>
<td>947.01</td>
<td>947.01</td>
<td>946.21</td>
</tr>
<tr>
<td></td>
<td>946.74</td>
<td>946.03</td>
<td>946.66</td>
<td>947.03</td>
<td>947.03</td>
<td>947.03</td>
<td>947.03</td>
<td>947.03</td>
</tr>
</tbody>
</table>

Concepts with prototypes
Concepts with naïve p.
Concepts with examplars

Concepts with prototypes
Concepts with naïve p.
Concepts with examplars

Concepts with prototypes
Concepts with naïve p.
Concepts with examplars

15
Multi-step planning
with effect-prototype-based concepts

Multi-step planning
with effect-prototype-based concepts
Multi-step planning
with effect-prototype-based concepts
Multi-step planning with naïve-prototype-based concepts
Multi-step planning
with exemplar-based concepts
Verb Concepts: Goal Specification

"iCub, do:

**********************************  - *****

Position along y

- Presence

Find most similar verb concept:

\[ f_{pro}^* = \arg \min_{f_{pro}} d_{EP}(f_{goal}, f_{pro}), \]

Find the behavior producing the verb concept best:

\[ b^* = \arg \max_{b} d_{EP}(\text{SVM}(e_{ok}, b), f_{pro}^*), \]

\[ d_{EP}: \text{Mahalanobis distance} \]
Mid-summary

• There are alternative ways for abstraction over behaviors/actions
• Prototype-based conceptualization based on effects is a good alternative
  – efficient planning
  – condensation
  – easy goal specification
  – Disadvantage: no information about the “how” part (not yet 😊).

Adjectives & Nouns based on Affordances & Visual Appearance

%15 Disappearable, %85 Pushable, %10 Knockable, %25 Grasppable

Therefore,
Short, thick, edgy
### Methodology: the Affordance Vector ($V_A$)

#### Probability of obtaining an effect from a behavior

<table>
<thead>
<tr>
<th>Behaviors vs Effects</th>
<th>PR</th>
<th>PL</th>
<th>PF</th>
<th>PB</th>
<th>TG</th>
<th>SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mov. Right</td>
<td>0.93</td>
<td>0.0</td>
<td>0.0</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Mov. Left</td>
<td>0.0</td>
<td>0.96</td>
<td>0.0</td>
<td>0.02</td>
<td>0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>Mov. Fwd.</td>
<td>0.0</td>
<td>0.0</td>
<td>0.89</td>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>Pulled</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.87</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Disappeared</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.09</td>
<td>0.0</td>
<td>0.03</td>
</tr>
<tr>
<td>Grasped</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.23</td>
<td>0.17</td>
</tr>
<tr>
<td>Knocked</td>
<td>0.03</td>
<td>0.02</td>
<td>0.08</td>
<td>0.0</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>No-change</td>
<td>0.04</td>
<td>0.02</td>
<td>0.03</td>
<td>0.0</td>
<td>0.64</td>
<td>0.47</td>
</tr>
</tbody>
</table>


---

#### Affordance Vector

48 x 124
Objects & labels

(a) cups
(b) boxes
(c) balls
(d) cylinders

(a) round
(b) edgy
(c) short
(d) tall
(e) thin
(f) thick
Predicted adjectives and nouns of novel objects.

<table>
<thead>
<tr>
<th>Object</th>
<th>$M_A^P$</th>
<th>Adjectives</th>
<th>$M_E^P$</th>
<th>$M_C^P$</th>
<th>$M_A^N$</th>
<th>Nouns</th>
<th>$M_E^N$</th>
<th>$M_C^N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_1$</td>
<td></td>
<td>edgy (54%)</td>
<td>edgy (89%)</td>
<td>edgy (60%)</td>
<td></td>
<td>box (74%)</td>
<td>box (97%)</td>
<td>box (56%)</td>
</tr>
<tr>
<td></td>
<td>short (97%)</td>
<td>short (55%)</td>
<td>short (80%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thin (59%)</td>
<td>thin (52%)</td>
<td>thin (52%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_2$</td>
<td>round (77%)</td>
<td>edgy (79%)</td>
<td>round (65%)</td>
<td></td>
<td>ball (83%)</td>
<td>ball (97%)</td>
<td>ball (80%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>short (77%)</td>
<td>short (58%)</td>
<td>short (68%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thin (89%)</td>
<td>thin 67%</td>
<td>thin (62%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_3$</td>
<td>edgy (63%)</td>
<td>edgy (64%)</td>
<td>edgy (60%)</td>
<td></td>
<td>cyl. (87%)</td>
<td>cyl. (95%)</td>
<td>cyl. (60%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>short (94%)</td>
<td>tall (67%)</td>
<td>tall (68%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thin (96%)</td>
<td>thin (84%)</td>
<td>thin (80%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_4$</td>
<td>round (84%)</td>
<td>round (77%)</td>
<td>round (75%)</td>
<td></td>
<td>box (94%)</td>
<td>cyl. (86%)</td>
<td>cyl. (52%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>short (98%)</td>
<td>short (68%)</td>
<td>short (71%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thick (91%)</td>
<td>thin (62%)</td>
<td>thick (51%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_5$</td>
<td>round (84%)</td>
<td>round (89%)</td>
<td>round (80%)</td>
<td></td>
<td>box (89%)</td>
<td>box (94%)</td>
<td>box (62%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>short (97%)</td>
<td>short (67%)</td>
<td>short (66%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thick (95%)</td>
<td>thick (58%)</td>
<td>thick (54%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_6$</td>
<td>edgy (84%)</td>
<td>edgy (79%)</td>
<td>edgy (75%)</td>
<td></td>
<td>cup (89%)</td>
<td>box (46%)</td>
<td>box (45%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>short (98%)</td>
<td>tall (55%)</td>
<td>short (65%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thin (92%)</td>
<td>thick (62%)</td>
<td>thick (52%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_7$</td>
<td>edgy (62%)</td>
<td>round (84%)</td>
<td>edgy (60%)</td>
<td></td>
<td>box (89%)</td>
<td>box (93%)</td>
<td>box (64%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>short (98%)</td>
<td>short (54%)</td>
<td>short (56%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thick (78%)</td>
<td>thick (68%)</td>
<td>thick (66%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_8$</td>
<td>round (72%)</td>
<td>edgy (89%)</td>
<td>round (62%)</td>
<td></td>
<td>cup (89%)</td>
<td>cup (98%)</td>
<td>cup (61%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>short (98%)</td>
<td>short (67%)</td>
<td>short (69%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>thick (79%)</td>
<td>thick (52%)</td>
<td>thick (53%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$M_A$: Learner from affordance vector
$M_E$: Learner from appearance
$M_C$: Learner from Appearance+Affordance
Predicted adjectives and nouns of novel objects from the KIT Dataset (Kasper et al., 2012).

<table>
<thead>
<tr>
<th>Object</th>
<th>Adjectives</th>
<th>Nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M^P_A$</td>
<td>$M^P_E$</td>
</tr>
<tr>
<td>$K_1$</td>
<td>round (60%)</td>
<td>edgy (92%)</td>
</tr>
<tr>
<td></td>
<td>tall (97%)</td>
<td>tall (100%)</td>
</tr>
<tr>
<td></td>
<td>thick (76%)</td>
<td>thick (96%)</td>
</tr>
<tr>
<td>$K_2$</td>
<td>round (55%)</td>
<td>edgy (90%)</td>
</tr>
<tr>
<td></td>
<td>tall (96%)</td>
<td>tall (98%)</td>
</tr>
<tr>
<td></td>
<td>thick (72%)</td>
<td>thick (91%)</td>
</tr>
<tr>
<td>$K_3$</td>
<td>edgy (55%)</td>
<td>edgy (92%)</td>
</tr>
<tr>
<td></td>
<td>tall (97%)</td>
<td>tall (95%)</td>
</tr>
<tr>
<td></td>
<td>thin (72%)</td>
<td>thin (93%)</td>
</tr>
<tr>
<td>$K_4$</td>
<td>round (58%)</td>
<td>edgy (76%)</td>
</tr>
<tr>
<td></td>
<td>tall (98%)</td>
<td>tall (100%)</td>
</tr>
<tr>
<td></td>
<td>thick (87%)</td>
<td>thick (86%)</td>
</tr>
<tr>
<td>$K_5$</td>
<td>round (55%)</td>
<td>edgy (76%)</td>
</tr>
<tr>
<td></td>
<td>tall (95%)</td>
<td>tall (98%)</td>
</tr>
<tr>
<td></td>
<td>thick (71%)</td>
<td>thick (94%)</td>
</tr>
<tr>
<td>$K_6$</td>
<td>edgy (59%)</td>
<td>edgy (83%)</td>
</tr>
<tr>
<td></td>
<td>tall (92%)</td>
<td>tall (96%)</td>
</tr>
<tr>
<td></td>
<td>thick (92%)</td>
<td>thick (90%)</td>
</tr>
</tbody>
</table>

$M^P_A$: Learner from affordance vector  
$M^P_E$: Learner from appearance  
$M^P_C$: Learner from Appearance+Affordance
Nouns vs. Adjectives

- Relevance of features to the category labels (acquired using ReliefF – Kononenko (1994))

Nouns prefer perceptual features whereas adjectives prefer affordance features.

Why?
Nouns vs. adjectives

• Psychology (Fernald, Thorpe, Marchman, 2009; Sandhofer, Smith, 2007):
  – Young children have more difficulty learning/interpreting noun modifying adjectives.

• Language (Sasson, 2011):
  – Adjectives are related to changes only in one/two dimensions whereas nouns depend on many dimensions in the feature space.
Conceptualization of Adjectives

Adjective prototypes obtained via learner with full affordance vector (\(V_{48}\))
(-): Highly confident that effect may not occur
(+): Highly confident that effect may occur
(*): Not confident about effect’s occurrence

<table>
<thead>
<tr>
<th>Adjective</th>
<th>TG</th>
<th>SG</th>
<th>PR</th>
<th>PL</th>
<th>PF</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edgy</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Round</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Short</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Tall</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Thin</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Thick</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

PR: Push Right, PL: Push Left, PF: Push Forward
PB: Pull, TG: Top Grasp, SG: Side Grasp

a: moved right  b: moved left
c: moved forward d: pulled
e: knocked       f: no change
g: Grasped       h: Disappeared

30
Co-learning nouns and adjectives

Orhan et al., “Co-learning nouns and adjectives”, submitted.
Mid-summary

• Nouns & Adjectives:
  – There is a functional/underlying difference between them
• This can shed some light to developmental psychologists & linguists

• Yuruten et al., "Learning Adjectives and Nouns from Affordances on the iCub Humanoid Robot ", SAB, 2012.
• Yuruten et al., “Learning of Adjectives and Nouns from Affordance and Appearance Features”, Adaptive Behavior, under revision.
Conclusion

• Theories on concepts from Psychology
• Hopefully, I have given some ideas:
  – for new experiments
  – explanations for existing ones
• There is still a lot to do regarding:
  – Verb Concepts
  – Adjectives
  – Nouns
Acknowledgments

Ilkay Atil, Asil Bozcuoglu, Yigit Caliskan, Nilgun Dag, Erol Sahin, Kadir Uyanik, Onur Yuruten.

Emergence of communication in RObots through Sensorimotor and Social Interaction

This project is also partially funded by TUBITAK through Projects 109E033 and 111E287.

More info available @ http://kovan.ceng.metu.edu.tr/