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

www.kovan.ceng.metu.edu.tr

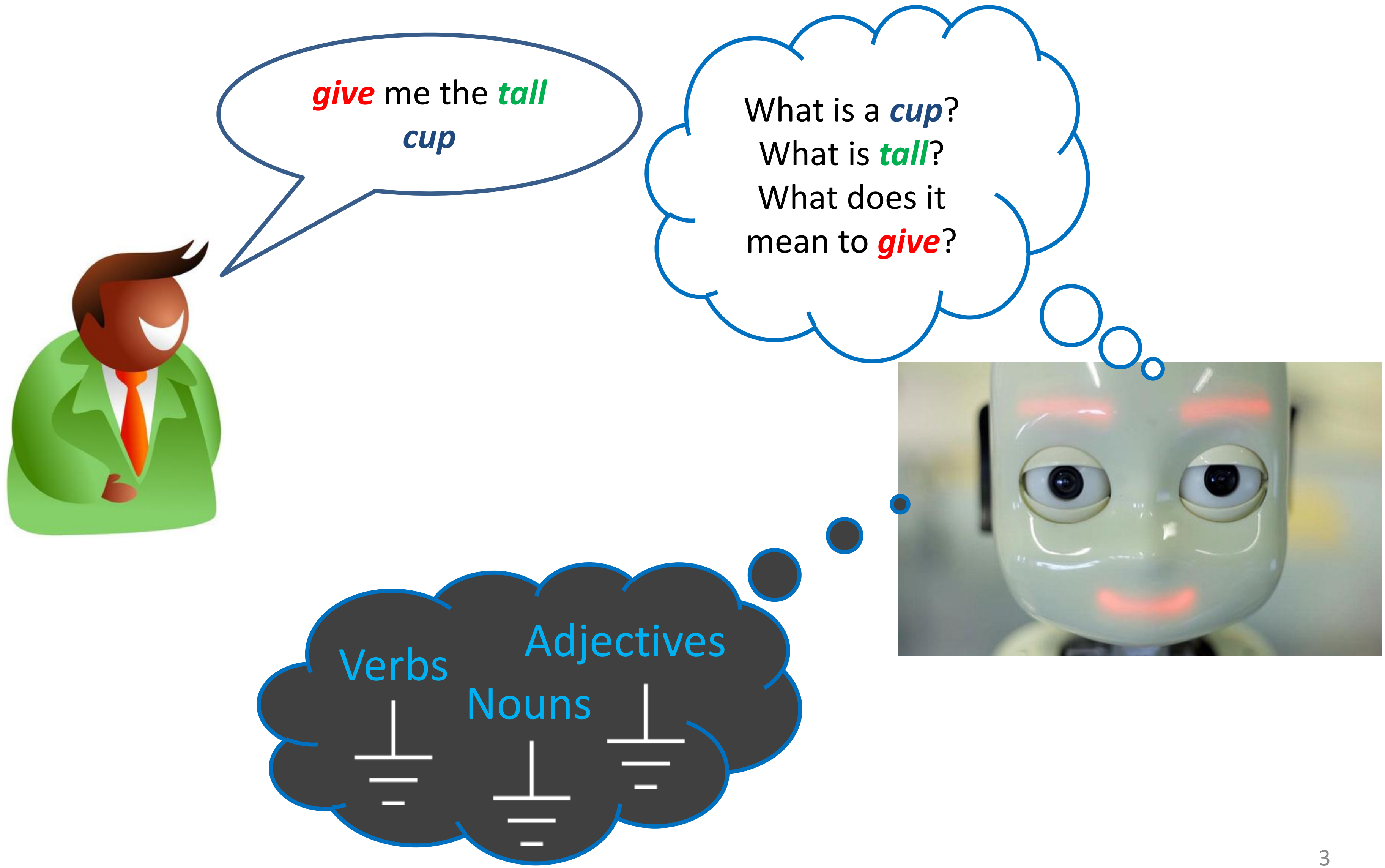
Vision

- Border Ownership
- Depth Prediction
- Feature Extraction
- Biometric Identification
- Image Retrieval

Cognitive Robotics

- Conceptualization & Affordances
- Multiple-Levels of Abstraction
- Context

Introduction: Problem



Introduction: Concepts

- Rule-Based

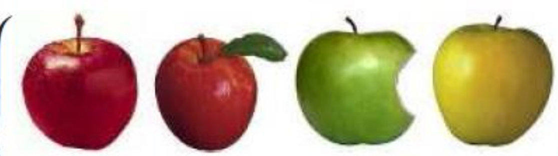
Apple = {color = “green” AND shape = “round”,…}

- Prototype-Based

OUR CHOICE

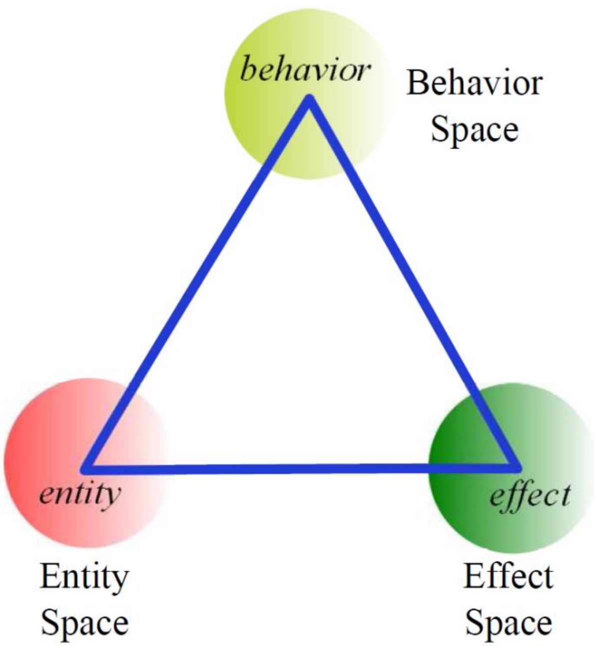
APPLE = {
colour RED, YELLOW or GREEN respectively
50%, 25% and 25% of the cases
shape 
…

- Exemplar-Based

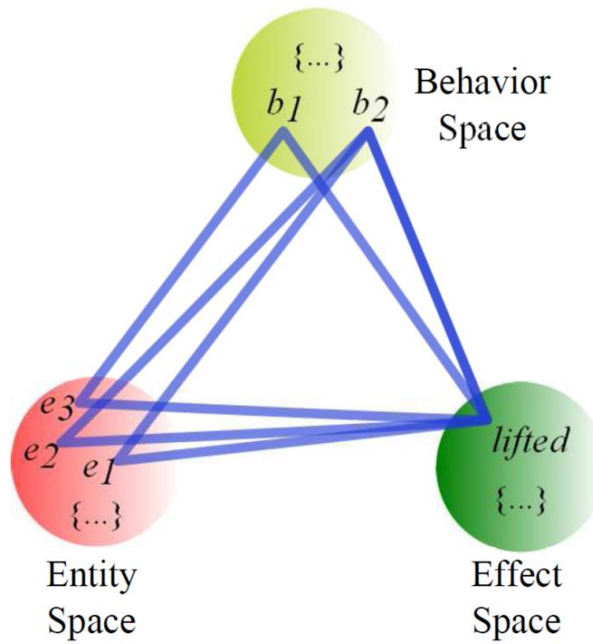
APPLE = {  … }

(Gabora et al., 2008; Kruschke, 2005; Rosch, 1973; Rouder & Ratcliff, 2006)

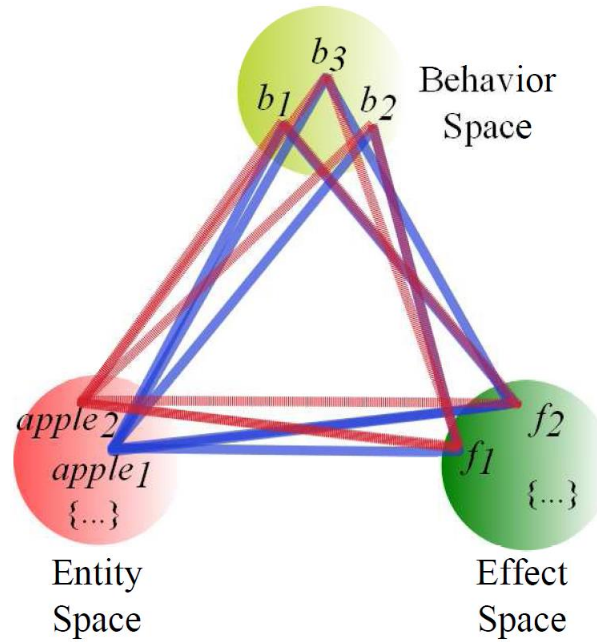
Introduction:
Affordances and Concepts



Affordance Formalization*



Adding Verb Concepts



Adding Noun/Adjective Concepts

* E.Sahin, M.Cakmak, M.R.Dogar and E.Ugur. *To Afford or Not to Afford: A new formalization of Affordances toward Affordance-based Robot Control. Adaptive Behavior, December 2007.*

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

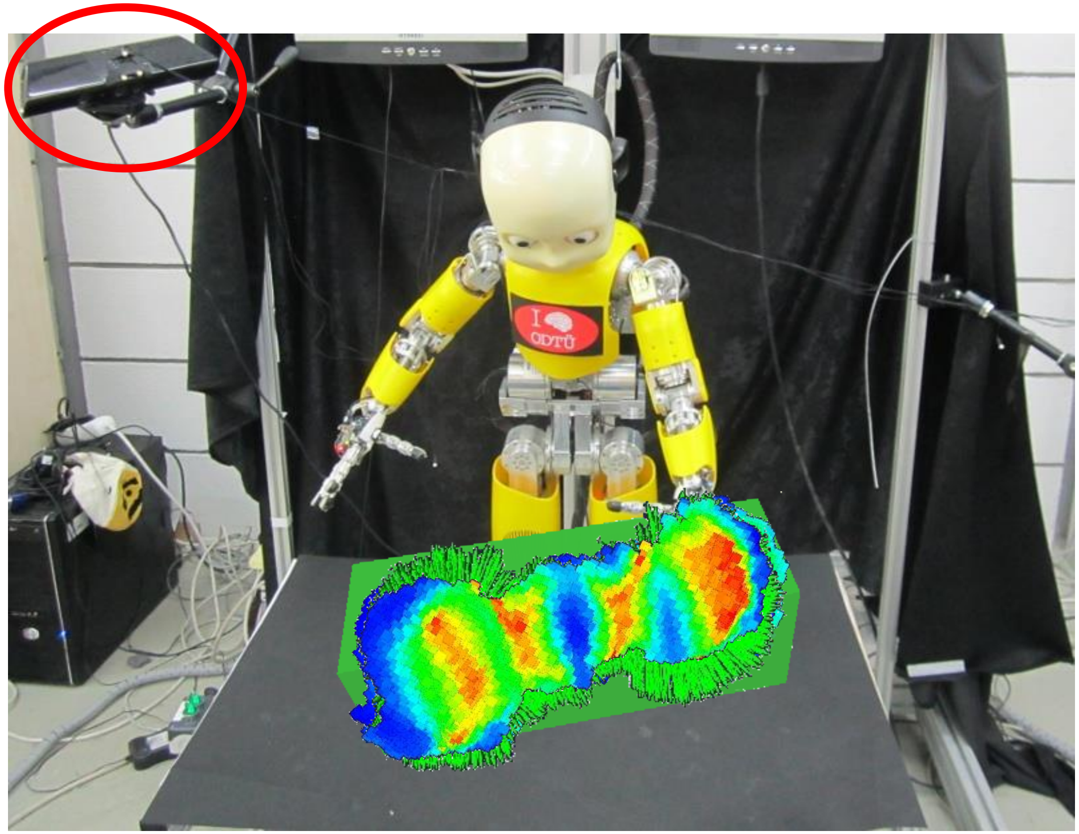
Noun/Adjective Concepts

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

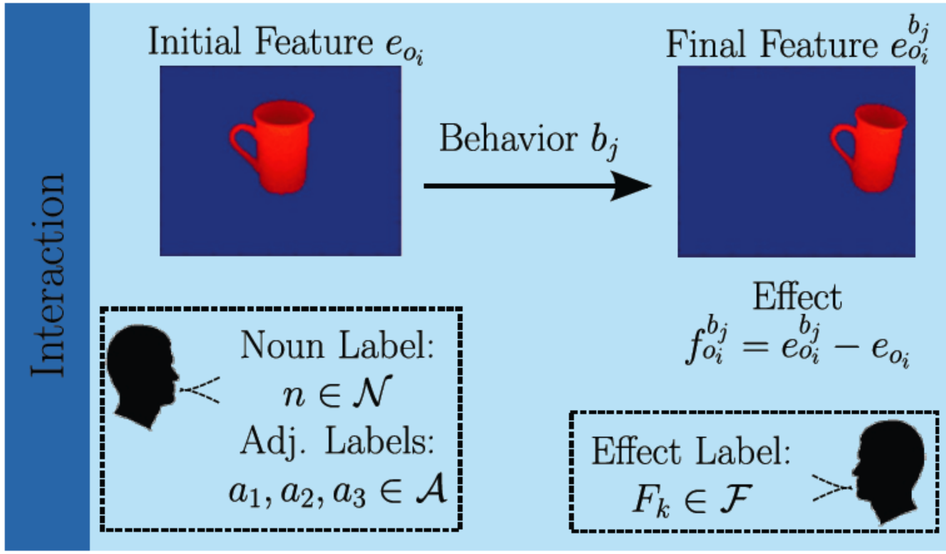


Features:

- 3D size, 3D position, surface normal histogram curvature histogram, presence

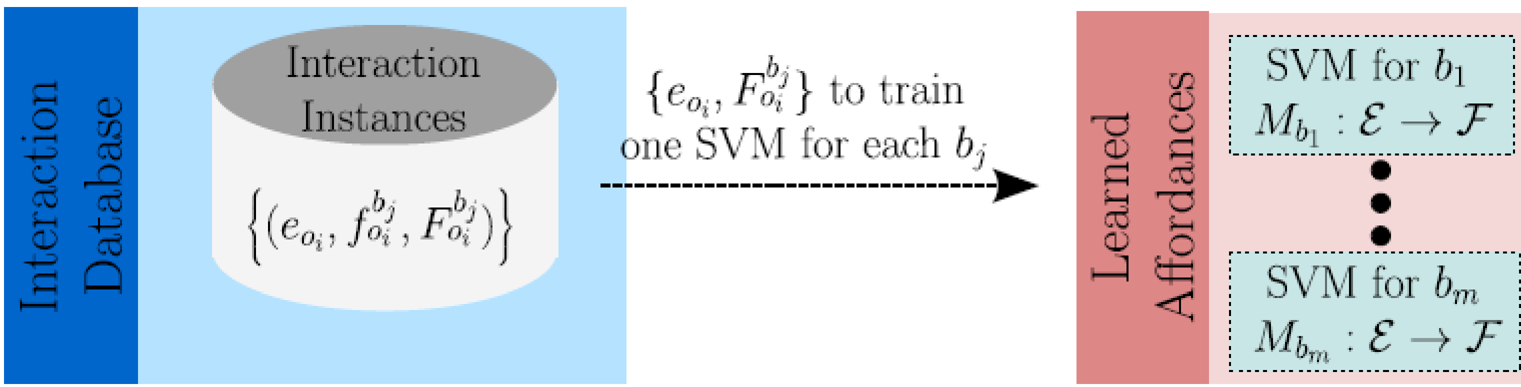
Nouns and adjectives

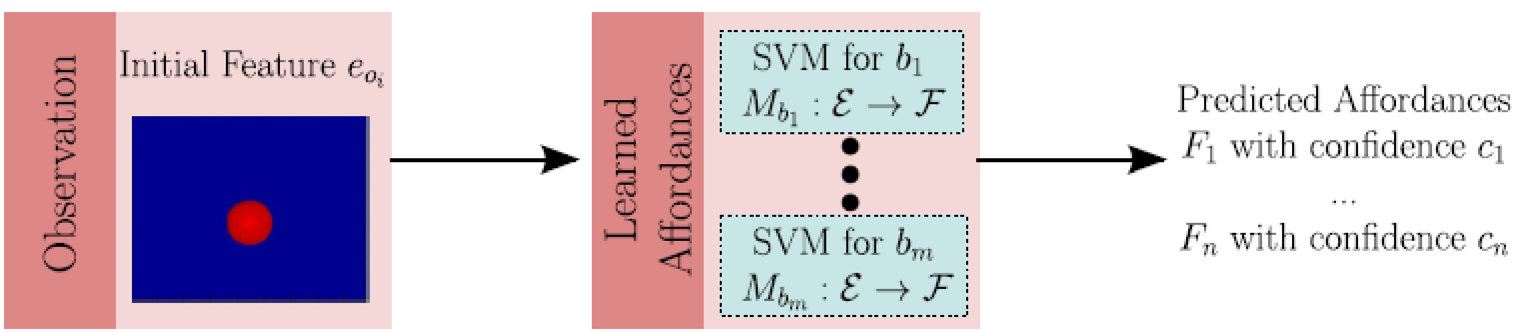
- **Cup**
- Box
- Cylinder
- Ball
- Short-tall
- Thin-thick
- Edgy-round



Verbs as effect labels

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





Experimental Setup:
Behaviors and Effects

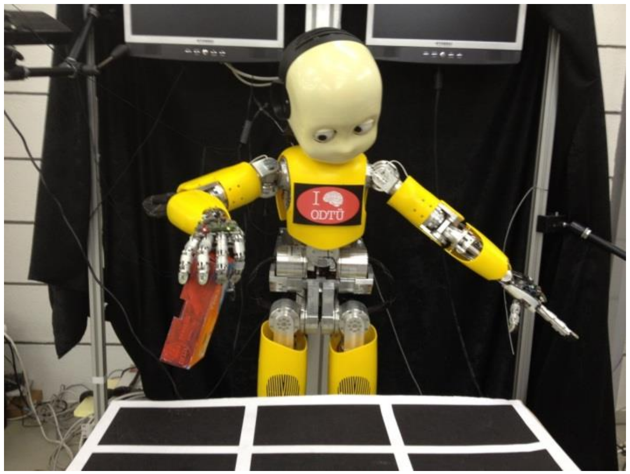


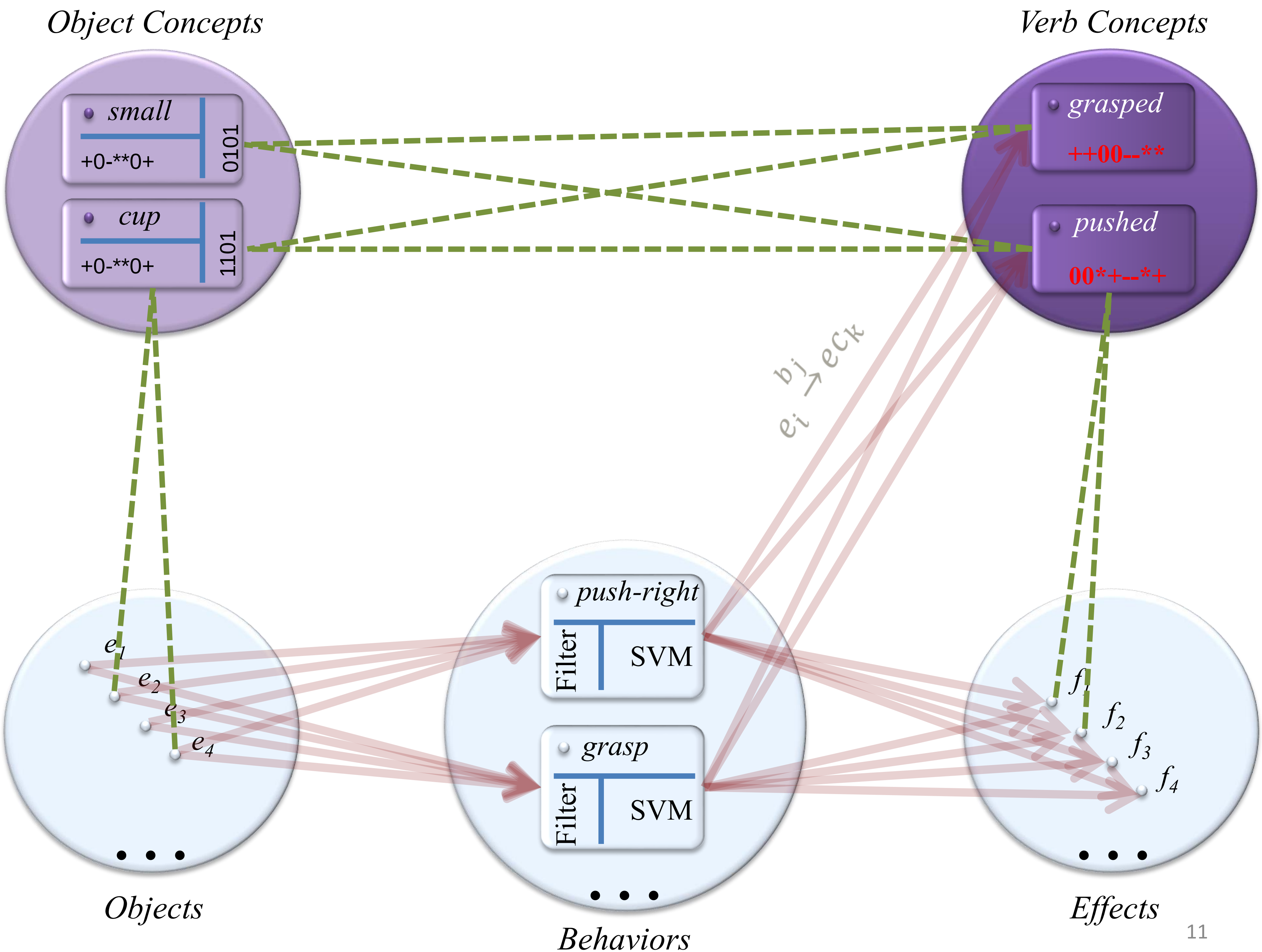
Figure 5. The objects interacted by the robot for learning.

Behaviors

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

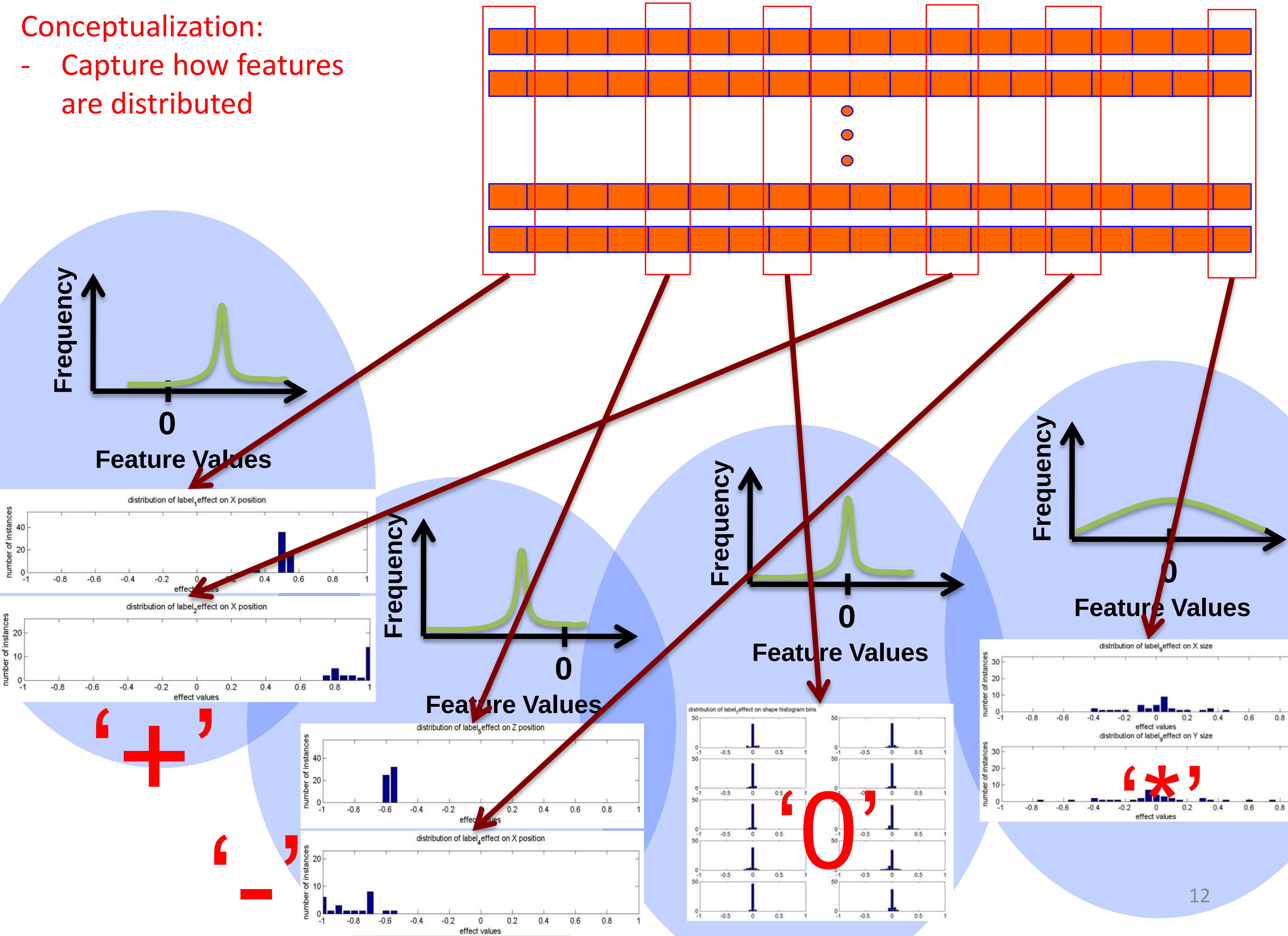
Effects

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


Object Concepts
Verb Concepts
small
+0-**0+
0101
cup
+0-**0+
1101
grasped
++00--**
pushed
00*+--*+
$e_i \xrightarrow{b_j} ec_k$
e_1
e_2
e_3
e_4
push-right
Filter
SVM
grasp
Filter
SVM
f_1
f_2
f_3
f_4
Objects
Behaviors
Effects

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

Eff. Cat. Name	Δ Azimuth Histograms	Δ Zenith Histograms	Δ Curvature Histograms	Δ Shape Index Histograms	Δ Position (x-y-z)	Δ Orient.	Δ Size (x-y-z)	Δ Object Presence
NC	*000000000 0000000000	0000000000 0000000*00	0000000000 00*0000000	0000000000 000000*000	000	0	000	0
MR	********** ******0*0*	********** *******000	0000000000 00********	*****0**** **0*0*0***	*+*	*	***	0
ML	******0000 00***0****	0**0**0000 **0000*000	0000000000 0*****0000	00*0**0*0* **0000*00*	0-0	*	000	0
MF	********** **********	******000* **000**000	0000000000 00****0***	***0*0**** **00000***	-00	*	000	0
P	********** **********	******000* **000**000	0000000000 00****0***	***0*0**** **00000***	+00	*	000	0
K	*0***00000 000000**0*	*000000000 0000000000	0000000000 000000000*	0000000000 0000000000	00-	*	00-	0
G	0000000000 0000000000	0000000000 0000000000	0000000000 0000000000	0000000000 0000000000	000	*	000	0
D	0000000000 0000000000	0000000000 0000000000	0000000000 0000000000	0000000000 0000000000	000	0	000	-

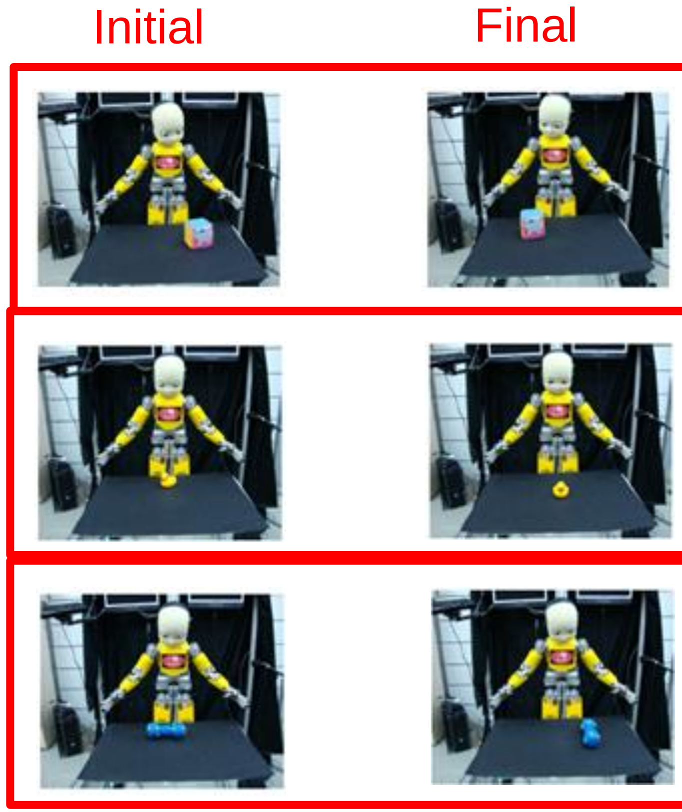
These prototypes also have mean and standard deviation values of the changes associated with each element.

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

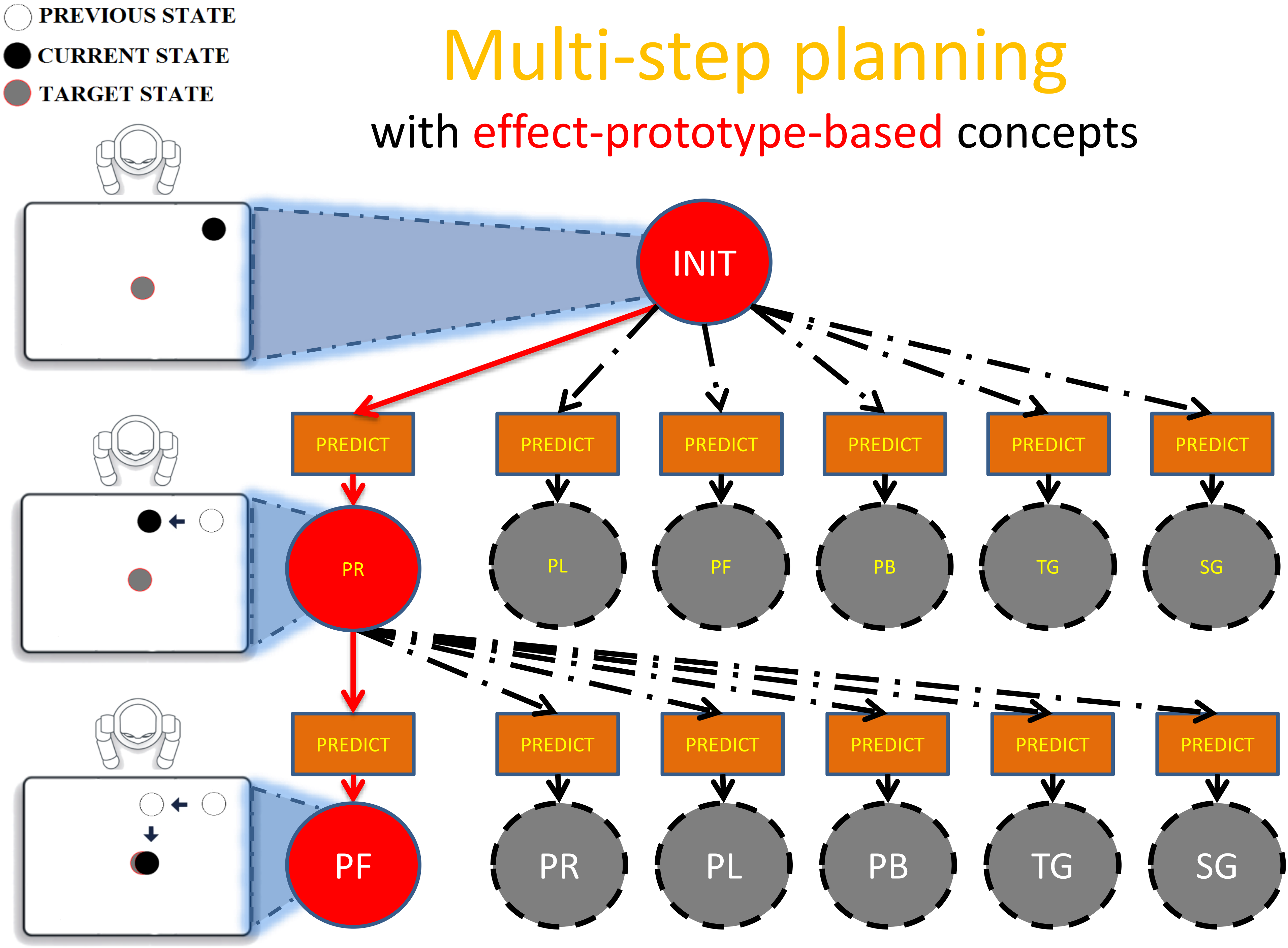
- Comparison using:
  - Recognizing an observed interaction
  - Planning

“what did I do?”



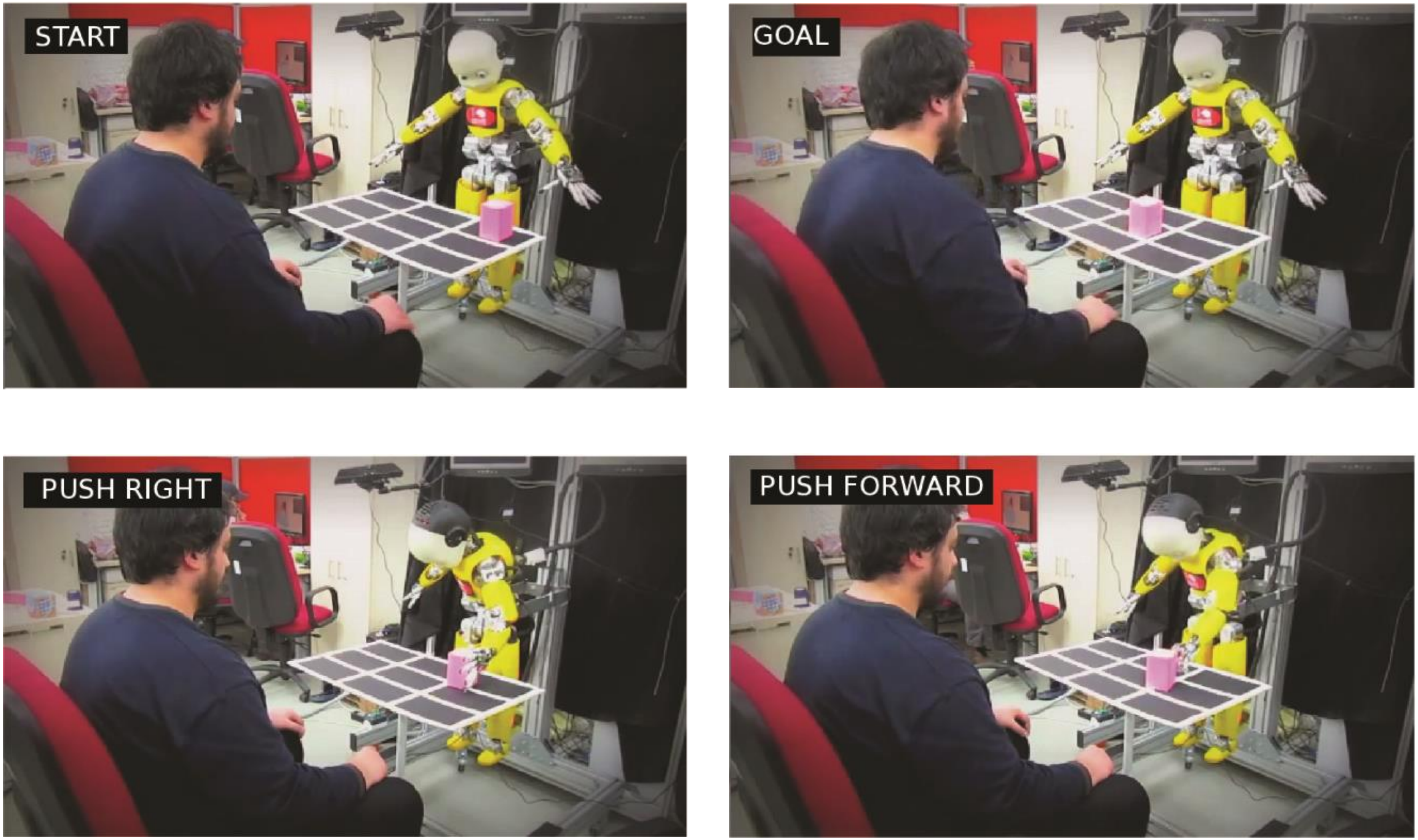
	No Change	Moved Right	Moved Left	Moved Forward	Pulled	Knocked	Grasped	Disappeared
Concepts with prototypes	390.81	146.24	372.24	389.21	215.56	215.50	392.11	410.31
Concepts with naïve p.	392.16	182.13	386.92	416.43	241.06	219.28	395.04	410.31
Concepts with examplars	237.01	236.89	237.42	237.24	237.42	237.42	237.25	236.84
Concepts with prototypes	731.36	494.18	416.42	340.71	393.76	358.06	738.04	790.41
Concepts with naïve p.	732.98	497.02	417.18	426.71	423.17	428.06	741.11	790.41
Concepts with examplars	789.45	789.08	789.83	789.49	789.83	789.83	789.54	788.84
Concepts with prototypes	925.41	577.51	267.45	328.75	354.85	354.74	928.16	947.51
Concepts with naïve p.	929.37	580.26	291.77	369.75	373.37	359.88	929.94	947.51
Concepts with examplars	946.74	946.42	947.03	946.66	947.03	947.03	947.01	946.21

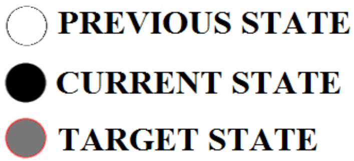
Multi-step planning
with effect-prototype-based concepts



Multi-step planning

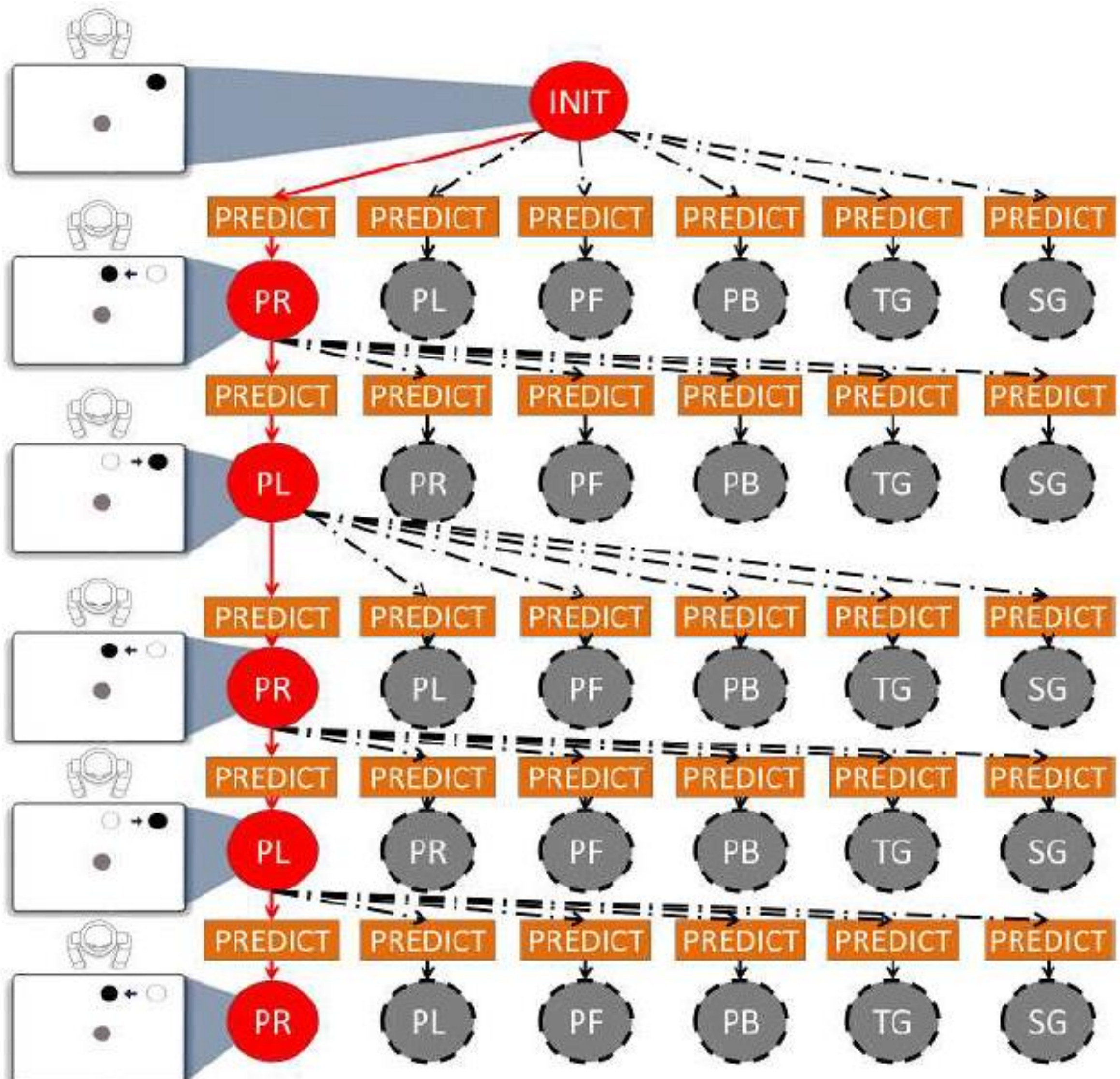
with effect-prototype-based concepts





Multi-step planning

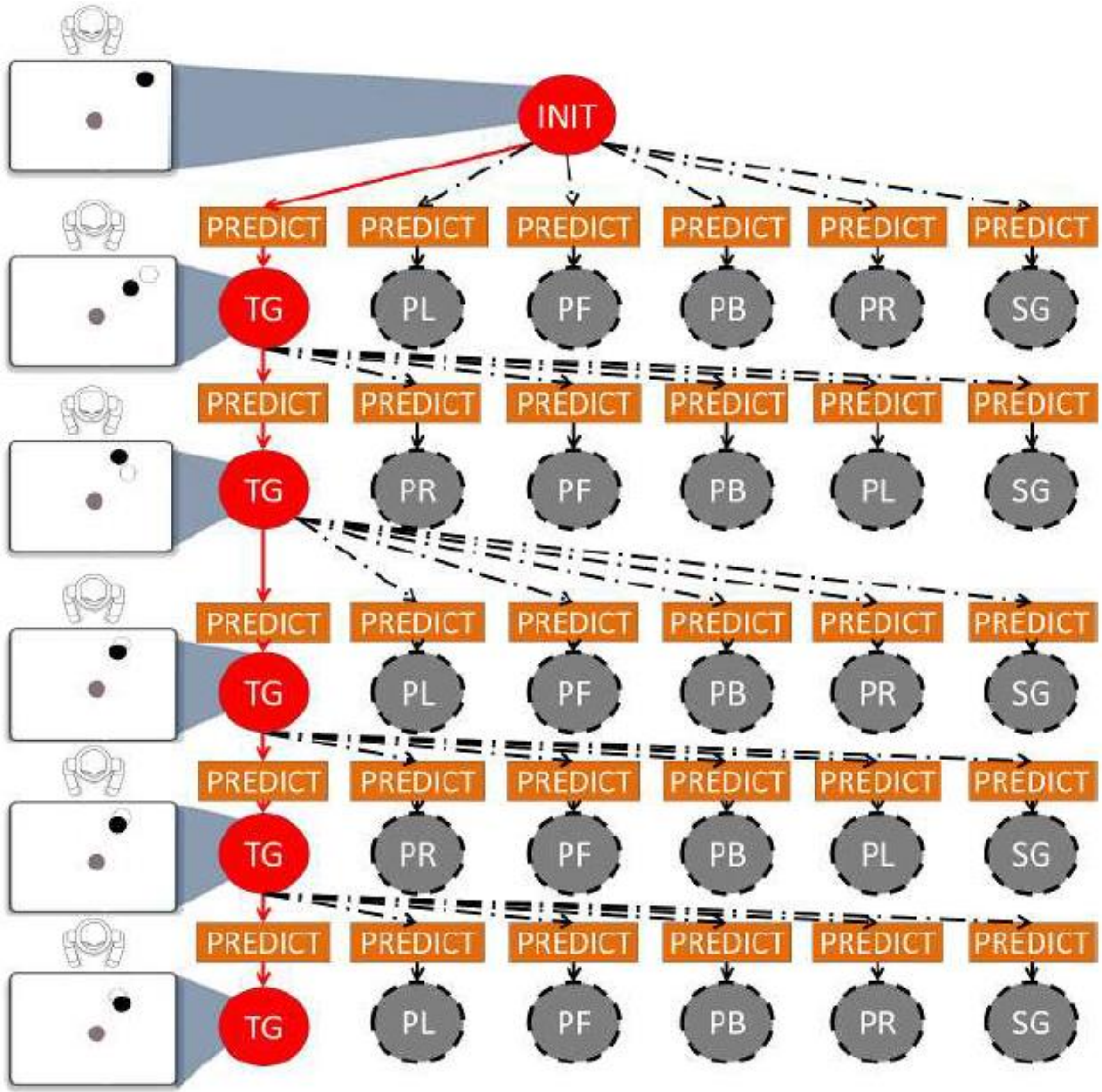
with naïve-prototype-based concepts



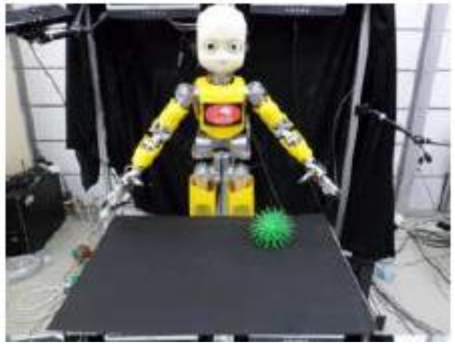
PREVIOUS STATE
CURRENT STATE
TARGET STATE

Multi-step planning

with exemplar-based concepts



Verb Concepts: Goal Specification



“iCub, do:
*********************************-*****0”

Position along y

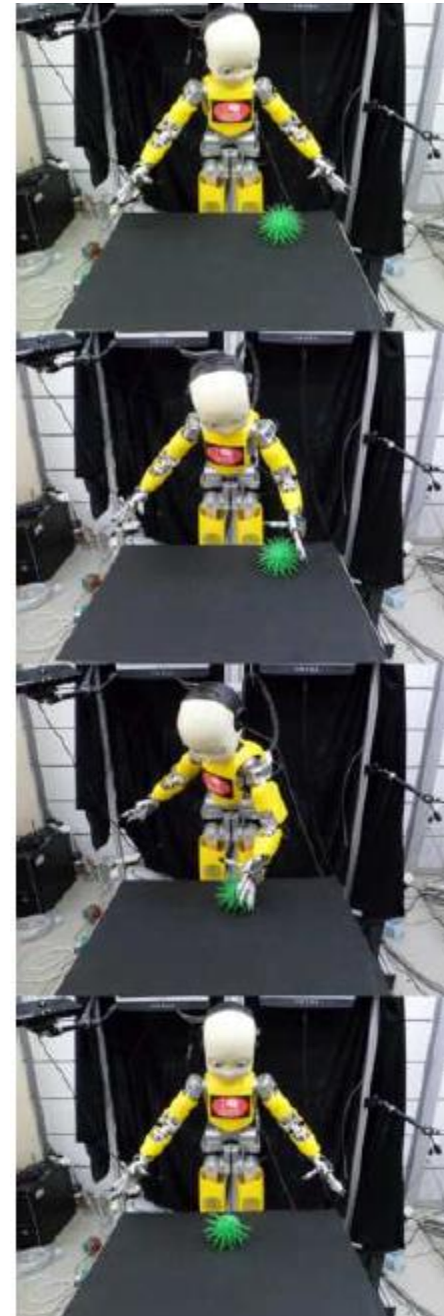
Presence

Find most similar verb concept:

$$f^*_{pro} = \arg\min_{f_{pro}} d_{EP}(f_{\text{goal}}, f_{pro}),$$

Find the behavior producing the verb concept best:

$$b^* = \arg\max_{b} d_{EP}(\text{SVM}(e_{o_k}, b), f^*_{pro}),$$



d_{EP}: Mahalanobis distance

Mid-summary

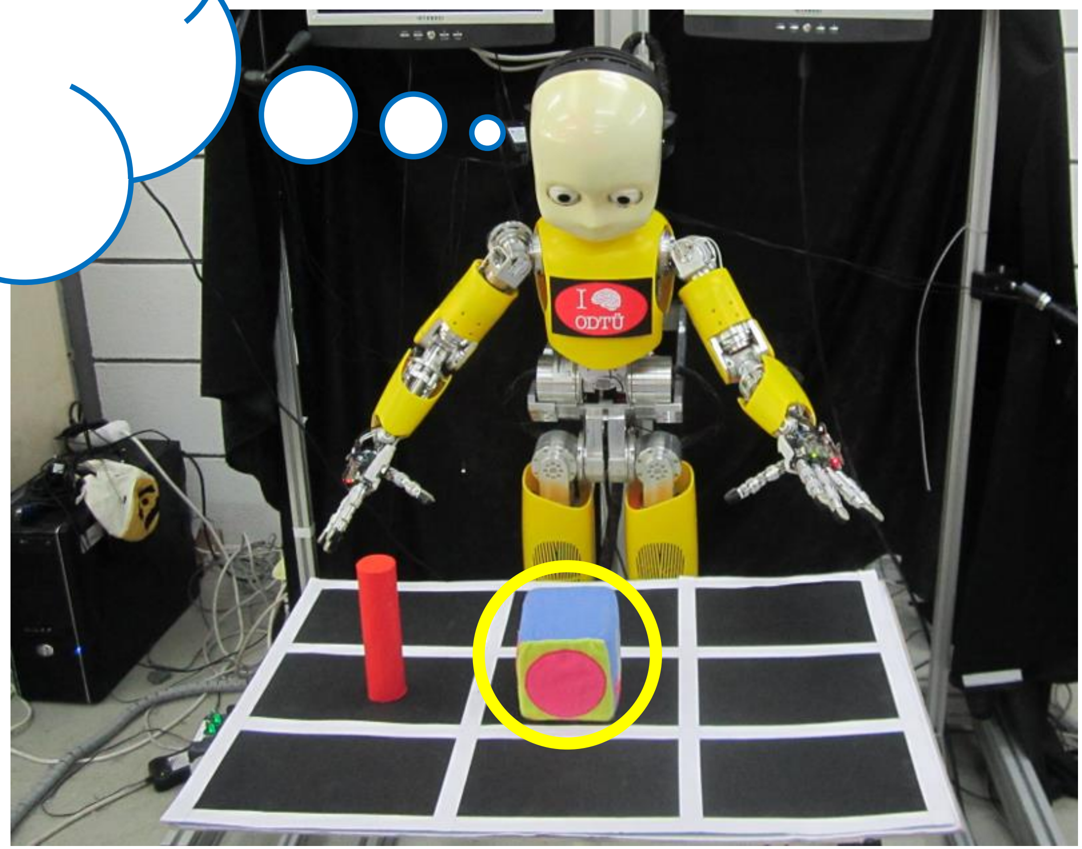
- There are alternative ways for abstraction over behaviors/actions
- Prototype-based conceptualization based on effects is a good alternative
  - efficient planning
  - condensation
  - easy goal specification
  - Disadvantage: no information about the "how" part (not yet ☺).

Kalkan et al., "Verb Concepts from Affordances", Interaction Studies Journal, in press.

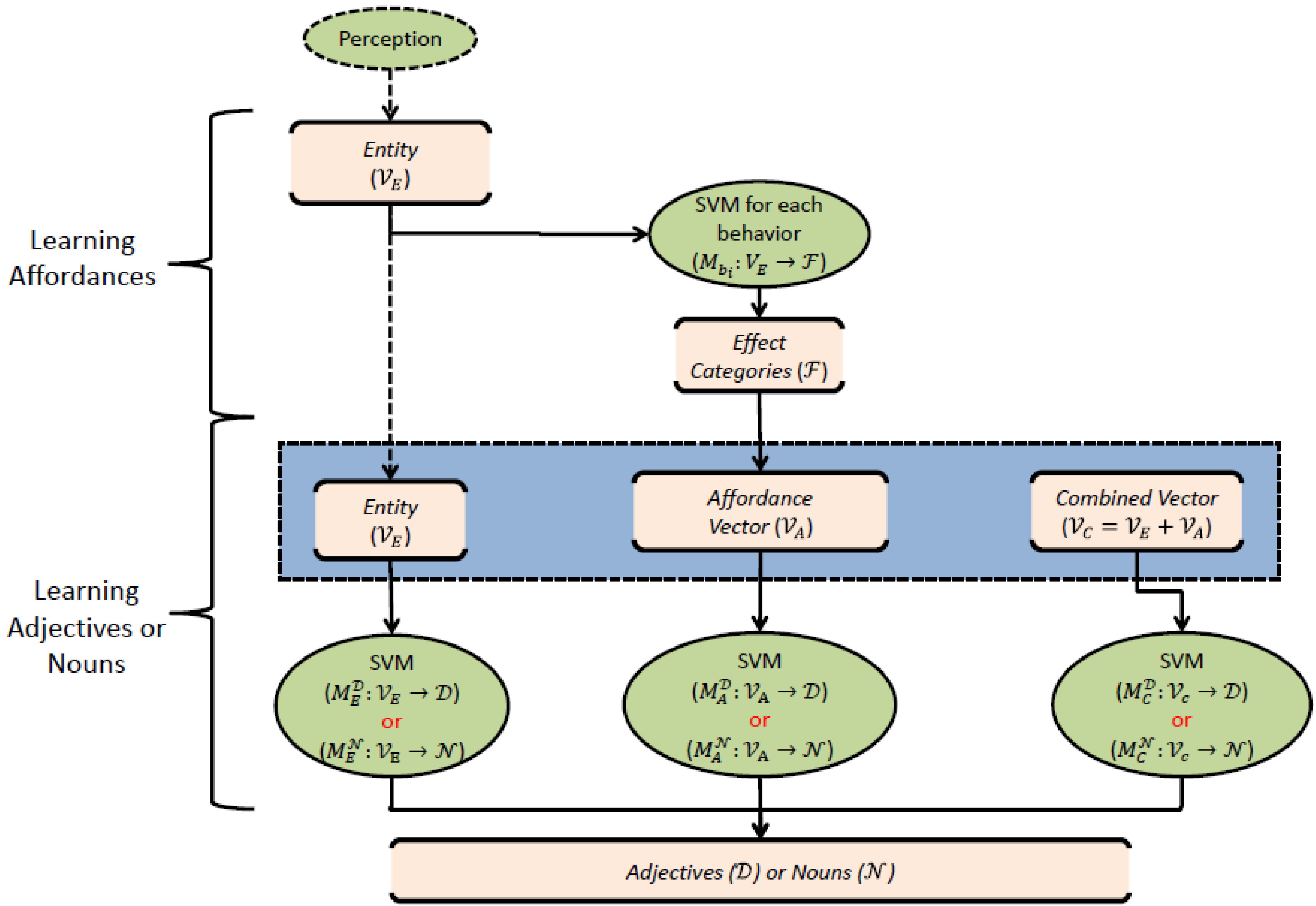
Adjectives & Nouns based on Affordances & Visual Appearance

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

Therefore,
Short, thick, edgy



Overview


Perception
Learning Affordances
Entity ($\mathcal{V}_E$)
SVM for each behavior ($M_{b_i}: \mathcal{V}_E \rightarrow \mathcal{F}$)
Effect Categories ($\mathcal{F}$)
Learning Adjectives or Nouns
Entity ($\mathcal{V}_E$)
Affordance Vector ($\mathcal{V}_A$)
Combined Vector ($\mathcal{V}_C = \mathcal{V}_E + \mathcal{V}_A$)
SVM ($M_E^{\mathcal{D}}: \mathcal{V}_E \rightarrow \mathcal{D}$) or ($M_E^{\mathcal{N}}: \mathcal{V}_E \rightarrow \mathcal{N}$)
SVM ($M_A^{\mathcal{D}}: \mathcal{V}_A \rightarrow \mathcal{D}$) or ($M_A^{\mathcal{N}}: \mathcal{V}_A \rightarrow \mathcal{N}$)
SVM ($M_C^{\mathcal{D}}: \mathcal{V}_c \rightarrow \mathcal{D}$) or ($M_C^{\mathcal{N}}: \mathcal{V}_c \rightarrow \mathcal{N}$)
Adjectives ($\mathcal{D}$) or Nouns ($\mathcal{N}$)

Methodology: the Affordance Vector (V_A)

Probability of obtaining an effect from a behavior

Behaviors vs **Effects**	**PR**	**PL**	**PF**	**PB**	**TG**	**SG**
Mov. Right	0.93	0.0	0.0	0.01	0.01	0.02
Mov. Left	0.0	0.96	0.0	0.02	0.03	0.15
Mov. Fwd.	0.0	0.0	0.89	0.01	0.01	0.04
Pulled	0.0	0.0	0.0	0.87	0.01	0.02
Disappeared	0.0	0.0	0.0	0.09	0.0	0.03
Grasped	0.0	0.0	0.0	0.0	0.23	0.17
Knocked	0.03	0.02	0.08	0.0	0.07	0.10
No-change	0.04	0.02	0.03	0.0	0.64	0.47

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

All Values

Affordance Vector

48 x 1

Objects & labels

(a) cups

(b) boxes

(c) balls

(d) cylinders

(a) round

(b) edgy

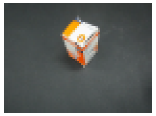
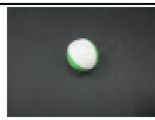
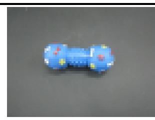
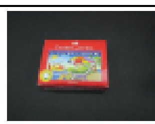
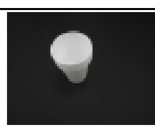
(c) short

(d) tall

(e) thin

(f) thick

Predicted adjectives and nouns of novel objects.

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

$\mathcal{M}_A$: Learner from affordance vector

$\mathcal{M}_E$: Learner from appearance

$\mathcal{M}_C$: Learner from Appearance+Affordance

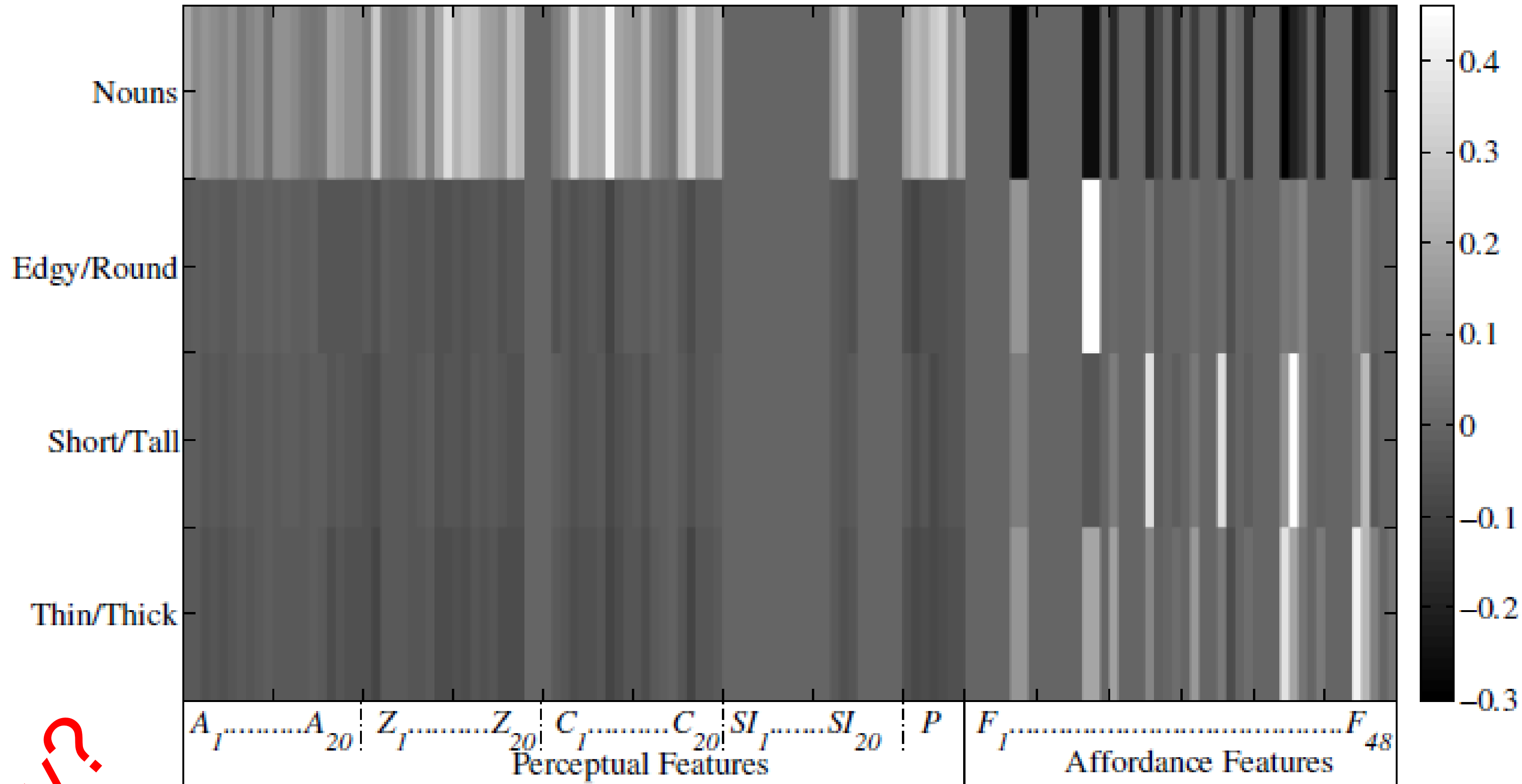
Predicted adjectives and nouns of novel objects from the KIT Dataset (Kasper et al., 2012).

Object		Adjectives $M_A^{\mathcal{D}}$	Adjectives $M_E^{\mathcal{D}}$	Adjectives $M_C^{\mathcal{D}}$	Nouns $M_A^{\mathcal{N}}$	Nouns $M_E^{\mathcal{N}}$	Nouns $M_C^{\mathcal{N}}$
K_1		round (60%) tall (97%) thick (76%)	edgy (92%) tall (100%) thick (96%)	round (60%) tall (80%) thin (52%)	cyl. (61%)	cyl. (98%)	cyl. (56%)
K_2		round (55%) tall (96%) thick (72%)	edgy (90%) tall (98%) thick (91%)	round (62%) tall (82%) thick (54%)	cyl. (56%)	cyl. (98%)	cyl. (58%)
K_3		edgy (55%) tall (97%) thin (72%)	edgy (92%) tall (95%) thin (93%)	edgy (92%) tall (79%) thin (81%)	box (58%)	box (97%)	box (59%)
K_4		round (58%) tall (98%) thick (87%)	edgy (%76) tall (100%) thick (86%)	edgy (82%) tall (83%) thick (70%)	cup (61%)	cup (96%)	cup (68%)
K_5		round (55%) tall (95%) thick (71%)	edgy (76%) tall (98%) thick (94%)	edgy (80%) tall (80%) thick (52%)	cup (56%)	cup (98%)	cup (56%)
K_6		edgy (59%) tall (92%) thick (92%)	edgy (83%) tall (96%) thick (90%)	edgy (62%) tall(78%) thick (52%)	box (56%)	box (99%)	box (62%)

$\mathcal{M}_A$: Learner from affordance vector
$\mathcal{M}_E$: Learner from appearance
$\mathcal{M}_C$: Learner from Appearance+Affordance

Nouns vs. Adjectives

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



Why?

- Nouns prefer perceptual features whereas adjectives prefer affordance features.

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

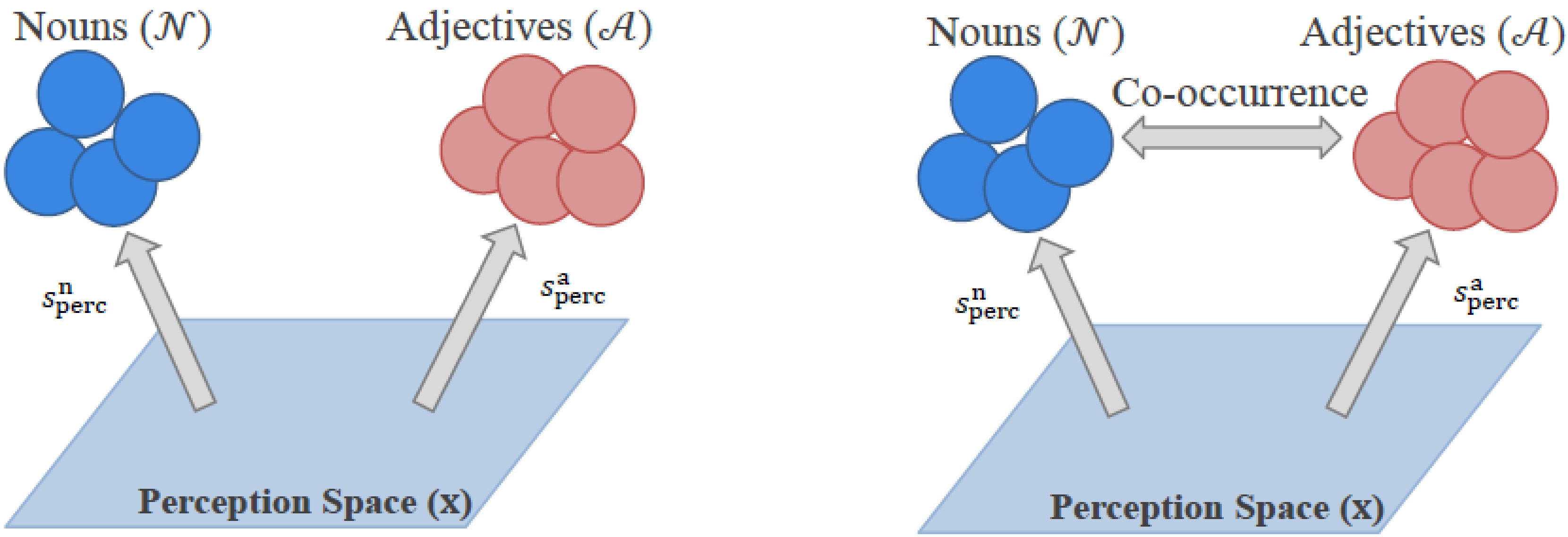
Adjective prototypes obtained via learner with full affordance vector (V_{48})
(-): Highly confident that effect may not occur
(+): Highly confident that effect may occur
(*): Not confident about effect's occurrence

Adjective	TG *abcdefgh*	SG *abcdefgh*	PR *abcdefgh*	PL *abcdefgh*	PF *abcdefgh*	PB *abcdefgh*
Edgy	−−−−−+−−	−−−−−**−	*−−−**−+	−*−−**−+	−−−***−+	−−−*++−+
Round	−−−−−**−	−−−−−+−−	*−−−+*−+	−*−−+*−+	−−−**+−*	−−−**+−*
Short	−−−−−**−	−−−−−+−−	+−−−**−+	−+−−**−+	−−−+**−+	−−−+*+−+
Tall	−−−−−**−	−−−−−**−	*−−−+*−+	−*−−+*−+	−−−*++−*	−−−*++−*
Thin	−−−−−**−	−−−−−**−	*−−−+*−+	−*−−+*−+	−−−*+*−+	−−−−++−+
Thick	−−−−−+−−	−−−−−**−	*−−−**−*	−*−−**−*	−−−**+−*	−−−+*+−*

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

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.
- Dag et al., *"Learning Affordances for Categorizing Objects and Their Properties"*, ICPR, 2010.

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.