Middle East Technical University
Department of Computer Engineering

Combining Different Knowledge-bases into a Single Partially-grounded Robotic Knowledge-base

İlker Bozcan, Sinan Kalkan

METU-CENG-TR-2017-02

March 2017

Department of Computer Engineering
Middle East Technical University
Dumlupınar Bulvarı, 06531, Ankara
TURKEY

Technical Report

This page contains a Turkish translation of the title and the abstract of the report. The report continues on the next page.

Farklı Bilgi-tabanlarını Kısmi Olarak Topraklanmış Bir Robot Bilgi-Tabanı Altında Birleştirme

İlker Bozcan, Sinan Kalkan

Bilgisayar Mühendisliği Bölümü
Ortadoğu Teknik Üniversitesi
Dumlupınar Bulvarı, 06531, Ankara
TÜRKİYE

Öz

Bu rapor, literatürde yer alan farklı bilgi-tabanlarını incelemekte ve bu bilgi-tabanlarının birleştirilmesi için bir yöntem önermektedir.

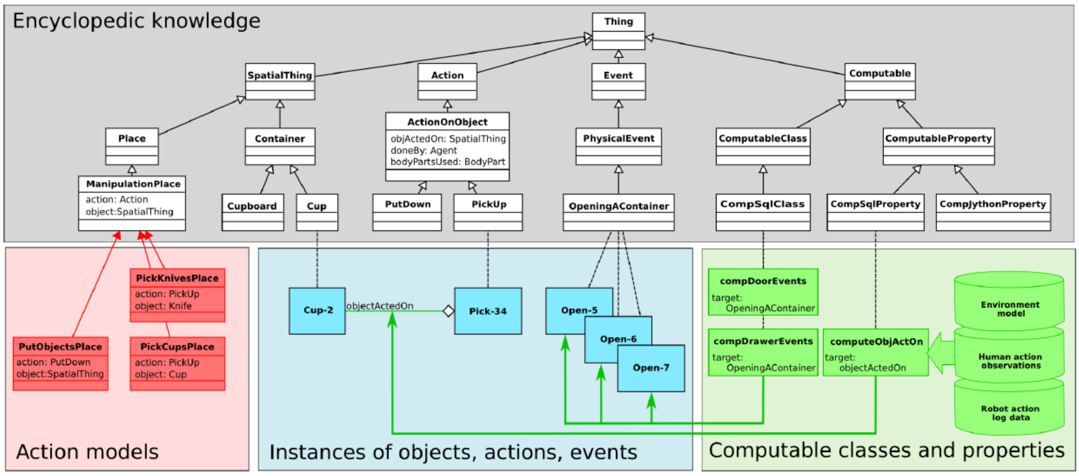


Figure 1: An overview of the KnowRob knowledge-base. (Figure source: [5])

Abstract

In this report, we review several robotic knowledge-bases openly available in the literature, and we propose to combine them into a single knowledge-base.

1 Knowledge-bases

Extending robot knowledge by using external knowledge sources (bases) is an important issue in robotics. There are several knowledge bases. All of these methods rely on two main paradigms: logic-based or graph-based. These knowledge representation methods have some advantages and disadvantages, and it is best if they are combined together.

1.1 KnowRob

KnowRob is the robot knowledge ontology proposed by Tenorth and Beetz [5] – see also Figure 1. In KnowRob, Knowledge is represented as formal logical statements using predefined templates. Widely used robotic concepts are ordered in hierarchical manner by using Web Ontology Language (OWL) syntax. There are variety of types of concepts which are related to robotics in the KnowRob ontology. For example:

- **Agent-Generic concepts** such as robot types (PR2,B21...), person.
- **Information Bearing concepts** such as semantic map of environment, floor plan.
- **Mathematical or Computational concepts** such as matrix, vectors, different types of algorithms.
- **Object type concepts** such as color, shape or intrinsic state (device is on/off etc.) of objects.
- **Spatial concepts** such as objects, object shapes, obstacles.
- **Temporal concepts** such as qualitative time of day (morning, afternoon).

A concept may have one or several parent concepts. Each concept may have several properties. For example, spatial object concepts may have spatial relations with other objects and event concepts may have `hasSubEvent` relation with some event concepts.

KnowRob allows making assertions/insertions using an OWL language:

```
owl_assert(manipActions,  onProperty,  actionType)
owl_assert(manipActions,  hasValue,    'ActionOnObject')
owl_assert(manipActions,  type,        'Restriction')

owl_assert(manipPosModel, type,        'ActionModel')
```

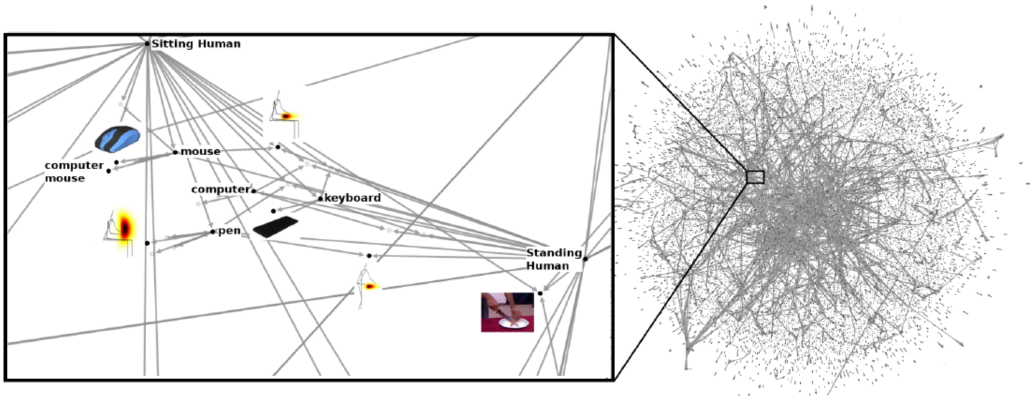


Figure 2: An illustration of the RoboBrain knowledge-base. (Figure source: [2])

and make queries in the knowledge-base via the same language, e.g.:

```
owl_query(?OVEN,     properPhysicalPartTypes, ?KNOB),
owl_query(?OVEN,     type, 'Oven'),
owl_query(?KNOB,     causes-Underspecified, ?HEATING),
owl_query(?HEATING,  postEvents, ?BOILING),
owl_query(?BOILING,  type, 'Boiling').
```

KnowRob has the advantage of compatibility with Open Cyc[1], a comprehensive ontology including everyday common sense knowledge.

1.2 RoboBrain

RoboBrain is knowledge engine introduced by Cornell and Stanford University in 2015 [2] – see Figure 2.

In contrast to KnowRob which represents knowledge in logical statements, RoboBrain represents knowledge as a directed graph. Each node represent concepts and edges represents relations between concepts. There can be several edges between two concepts. Knowledge graph can be extended by external sources if edges from the edge set of RoboBrain is used. Robot Query Language (RQL) is used for reasoning in the graph.

RQL can be used as interface between the agent and knowledge graph. The agent can query possible affordances of perceived object , get trajectory to apply desired affordance or get parameters of given trajectories etc. For instance, consider query q as "*affordances n:= fetch({name:n})→ ['HasAffordance'] → (v{src:'Affordance'})*". q returns possible object affordances for each object which is intereseted in the environment. Now, consider query q_2 as *trajectories a:= fetch({handle : a}) → ['HasParameters'] → (v{src: 'Affordance', type: 'Trajectory' })* . q_2 returns motion trajectories for each affordances.

1.3 Knowledge Base(KB) using Markov Logic Network(MLN)

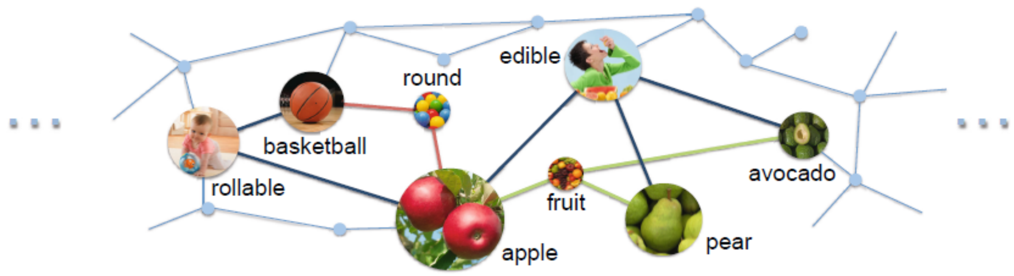
A probabilistic graphical model for representing information between object-action concepts was introduced by Yuke Zhu, Alireza Fathi and Li Fei-Fei from CS Department of Stanford University [7] – see Figure 3. In this work, knowledge is represented as a Markov Logic Network, composed of Markov Random Field and First Order Logic.

In this model, as in RoboBrain, nodes correspond to concepts and edges correspond to relations. There can be only one undirected edge between two nodes unlike in RoboBrain.

Concepts have four categories: instances (objects), categories, affordances and attributes (weight, size or visual attributes).

There are two main sources used to populate the knowledge graph: web and dataset. Attribute concepts are collected from Amazon, eBay. Categorical concepts are collected from FreeBase, and

[1] http://www.opencyc.org/



(a)

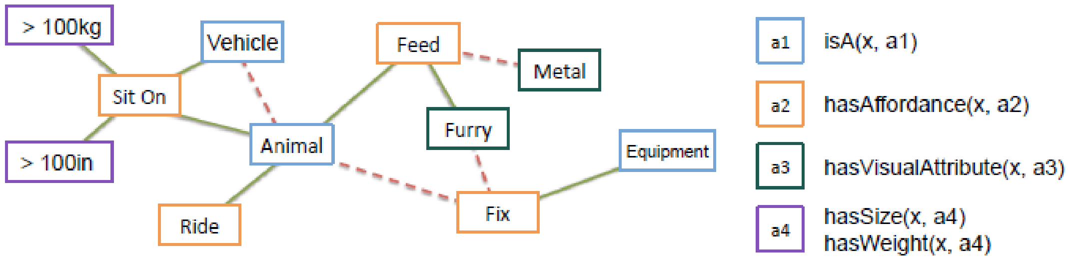


Figure 3: An illustration of the probabilistic graphical knowledge-base. (Figure source: [7])

WordNet. Affordance concepts are collected from Google Ngram or manually labeled. The dataset is used for creating visual attributes of objects, infer the affordances of objects.

Using RoboBrain, affordance prediction of a novel object, estimating human pose for an affordance and question answering by using knowledge graph is applicable.

1.4 ConceptNet

ConceptNet [4] is a free semantic network originated from Open Mind Common Sense project [3] – see Figure **??** for a snapshot. It includes several kinds of relations between concepts that might be helpful for robotic research. There are denser connection between concepts than projects mentioned above.

ConceptNet 5 is the most up-to-date version. It contains more than 2 million concepts and approximately 28 million total edges. ConceptNet provides weights for each edge determined by volunteer participants and knowledge sources on the web. Edges are helpful while determining semantic similarity between concepts.

ConceptNet is also multilingual, it contains 10 core languages which is supported by ConceptNet 5. They are English, French, Italian, German, Spanish, Russian, Portuguese, Japanese, Dutch, Chinese. Furthermore, there are 68 languages that has vocabulary with more than 10.000 words.

There can be variety of types of edges in a ConceptNet graph. Each edge between two concepts indicates a relation between them. For instance, **relatedTo** is the most general relation between two concepts. It indicates positive relation between concepts. **isA** relation indicates subtype-supertype relation between two concepts. Full list of relations can be found in [4].

1.5 Probabilistic Concept Web

The probabilistic concept web [1] is a probabilistic graphical representation of robot knowledge based on Markov Random Fields. As in graph representation of knowledge sources, concepts are nodes of the graph and edges are relations between nodes – see Figure 5 for a snapshot.

Each concept is represented by their prototypes that is extracted from training set by using proposed method in [1]. Edge weights between concepts are determined by co-occurrences of concepts in robot-object interaction.

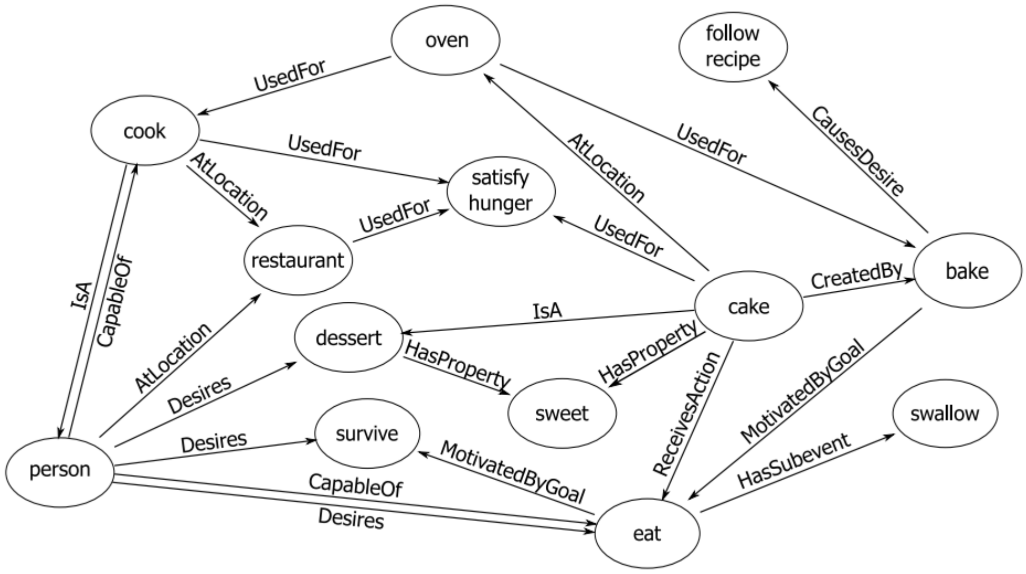


Figure 4: A portion of a cluster with related concepts from ConceptNet (Figure source: [4])

There are three types of **grounded concepts** according to robot sensor space in the probabilistic concept web:

- Noun concepts: {ball,box,cup,cylinder,plate,tool}.
- Adjective concepts: {edgy,round,noisy,silent,tall,short,hard,soft,thin,thick}
- Verb concepts: {drop, grasp, moveForward, moveBackward, moveLeft, moveRight, pushBackward, pushForward, pushLeft, pushRight, shake, throw}

In original work [1], three scenarios are used to test the concept web. Firstly, iCub perceives the object visually and guesses about its properties. Using Concept Web convergence, properties (adjective concepts such as hard, round etc.) of objects are refined and which behaviors (verb concepts shake, drop etc.) are applicable is determined.

Secondly, iCub is said to apply some behavior to given object. In this scenario, object concepts and verb concepts are activated. After convergence in the MRF graph, hidden concepts that indicate object properties are activated. Haptic and audio related properties of objects are determined without touching the object using by MRF graph.

In last scenario, iCub is said to apply some behavior to multiple object on the environment. Using Concept Web, iCub decides which object is the best one to apply desired behavior.

1.6 A Comparison of the Knowledge-bases

As stated in Table 1, Concept Web, RoboBrain and KnowRob are robotic projects. They try to represent knowledge in order for robots to manipulate objects better. KB using MLN is a knowledge representation model for extracting knowledge from images. ConceptNet is knowledge representation for general purpose in AI and Natural Language Processing.

Only in concept web, knowledge in the graph is fully grounded in robot's sensor level. Therefore, each concept has semantic according to robots sensor stimuli. In RoboBrain and KnowRob, concepts are also grounded but they include concepts that are extracted from web without grounding. Other two works are not concerned with grounding of concepts.

Concept Web includes concepts that are created by using haptic, audio and visual sensors. Therefore, concept semantic has created according to the richest diversity of sensor data. ConceptNet does not include physical semantic for concepts. Each concept may interpreted by its relations with other concepts instead of sensor data. Other works, concept grounding is made

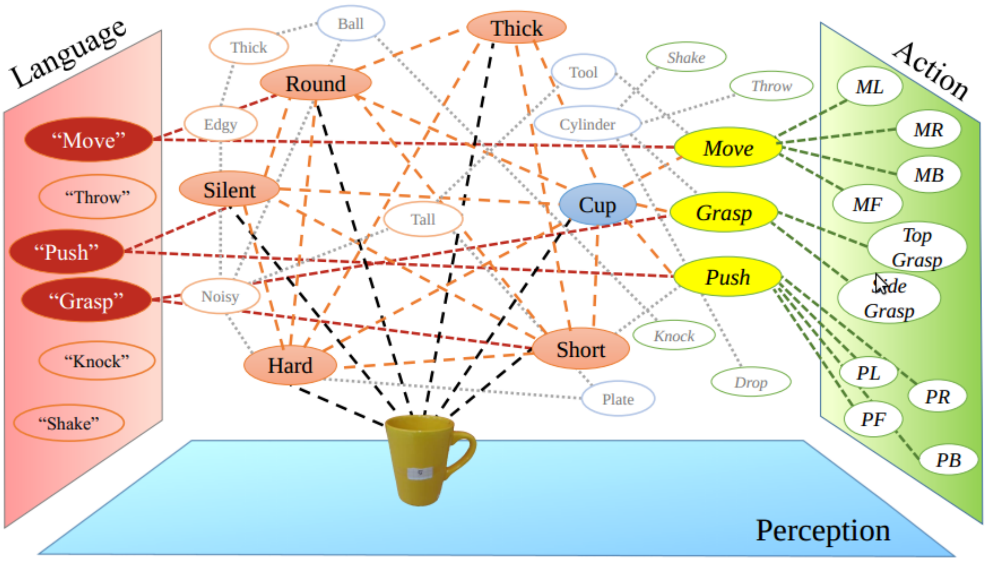


Figure 5: Schematic presentation of Scenario 1. iCub is presented with a cup, allowed to interact with it freely, and expected to predict the type and properties of the object, as well as what kind of behaviors can be applied on this object. The converged concept web is depicted. The action space and verb concepts are contoured with green, whereas blue and orange colors represent the noun and adjective categories for the object, respectively. The gray and smaller fonts show inactive concepts in the web, while bigger fonts and colored nodes represent activated concepts. There are other concepts and connections that are not shown for clarity. ML: Move-Left, MR: Move-Right, MB: Move-Backward, MF: Move-Forward, PL: Push-Left, PR: Push-Right, PB: Push-Backward, PF: Push-Forward. (Figure source: [1])

using by visual data including images, 3D points, videos etc. KB using MLN is a web as a source for extracting size and weight of objects also.

Concept Web and KB using MLN are based on Markov Random Field Theory (MRF). In MRF, relations between nodes have weight represented as conditional probability of activating each other. This theory allows probabilistic inference of knowledge instead of deterministic ones. In ConceptNet, knowledge is represented by directed graph and edges have weight. However, unlike in Concept Web and KB using MLN, weights are determined hand-coded by using web source and volunteers. KnowRob represents knowledge as formal statements using predefined templates.

Knowledge acquisition is made by robot in Concept Web, RoboBrain and Knowrob. However, RoboBrain and KnowRob use web as knowledge source also. KB using MLN uses Stanford 40 Actions dataset[6] and web. ConceptNet knowledge is crawled by using web and participating volunteers.

There is no interface between knowledge representation model and agent in Concept Web. KB using MLN, RoboBrain and KnowRob use their own query languages as interface and ConceptNet has Python API for accessing knowledge graph.

2 Merging Concept Web, ConceptNet and KnowRob

We transfered concept web [1] concepts into KnowRob ontology. Affordances for objects in KnowRob are not defined. Therefore, by using concept web, we assign affordances for common noun concepts. In addition, we assign probability of having some property given in adjectives to each noun concept.

Using ConceptNet 5, we include several super concept and subconcept of given concepts. We

Table 1: Comparison of knowledge representation models

	Probabilistic Concept Web [1]	**KB using MLN** [7]	**ConceptNet** [4]	**RoboBrain** [2]	**KnowRob** [5]
Field	Robotic	Image	AI	Robotic	Robotic
Grounded	Yes	No	No	Yes (Partially)	Yes (Partially)
Perception	Visual Haptic Audio	Image Weight&Size	No	Grasping features 3D clouds images, videos etc.	3D clouds
Representation	Graph (MRF)	Graph (MLN)	Directed graph	Directed graph	Formal statements
Knowledge Acquisition	Robot	Web, Dataset	Volunteers, Web	Web/Robot	Web (Cyc) Hand-coded Robot
Interface	None	Queries	Web/API	Queries (RQL)	Queries

can extract information about where an object can store and which behaviors are applicable for this object. Attributes of *ball, box* and *cup* concepts in Concept Web are enriched by using ConceptNet 5.

Final properties of merged concepts are here, only some examples of common concepts in three knowledge representation models are shown below. Source of the knowledge is provided in parentheses:

- **ball**
  - is subconcept of *PortableObjects* (KnowRob)
  - may have spatial properties with objects such as being near of some object etc. (KnowRob)
  - has affordance of move, push, shake, throw (Concept Web)
  - has probability of some adjective properties: soft, noisy, short, thick, round (Concept Web)
  - can be stored in a toybox (ConceptNet)
  - have subtypes such as soccer ball, beach ball (ConceptNet)
  - is used for throwing, bouncing, playing a game (ConceptNet)
- **box**
  - is subconcept of *SpatialThingTypeByShape* (KnowRob)
  - may have spatial properties with objects such as being near of some object etc. (KnowRob)
  - has affordance of drop, grasp, move, push, shake, throw (Concept Web)
  - has probability of some adjective properties: soft, noisy, short, thick, edgy (Concept Web)
  - can be stored in a any garage, ballpark (ConceptNet)
  - have subtypes such as ballot bax (ConceptNet)
  - is used for storing something in (ConceptNet)
- **cup**
  - is subconcept of *DrinkingVessel* (KnowRob)
  - may have spatial properties with objects such as being near of some object etc. (KnowRob)
  - is stored in some cup board (KnowRob)
  - has subconcept *mug* (KnowRob)
  - has affordance of grasp, move, push (Concept Web)

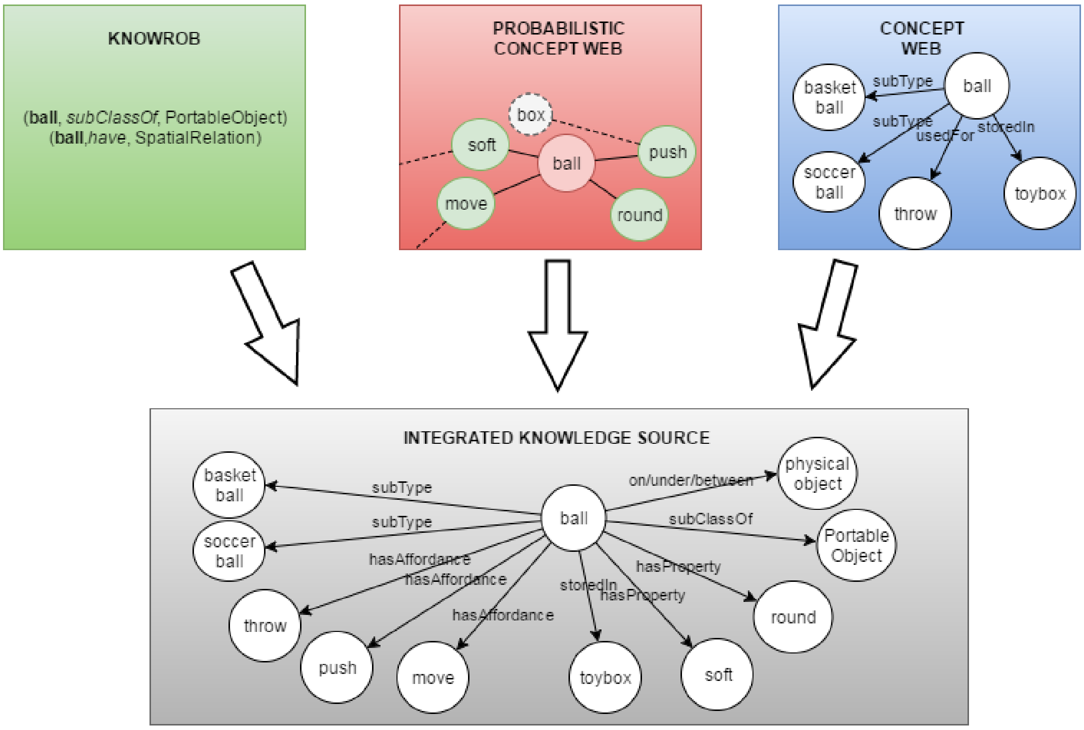


Figure 6: In this figure, integration of three knowledge sources is shown. As an example, attributes of "cup" concepts in different knowledge sources are merged and new "cup" concept is created in new knowledge source. KnowRob [5] is indicated in green rectangle. It includes formal statements related to properties of "cup" concept. From KnowRob, categorical and spatial properties of "cup" concept are extracted. Probabilistic Concept Web [1] is shown in red rectangle. "cup" concept is shown in red node that is interested. Green nodes that are connected with solid line are active concepts that are connected to "cup" concept. Dashed lines indicate inactive connections. From Concept Web, affordances and properties of "cup" concept are got. Concept Net[4] is shown in blue rectangle. Relations between "cup" concept and other concepts shown with directed arrow. From Concept Net, categorical properties, affordances and storage information are got. Integration of three knowledge sources is shown in gray rectangle. New "cup" concept is created with relations which are combined from these sources.)

    - has probability of some adjective properties: hard, silent, short, thick, round (Concept Web)
    - can be stored in a table, shelf (ConceptNet)
    - have subtypes such as champagne cup (ConceptNet)
    - is used for drinking (ConceptNet)

- **cylinder**
    - is subconcept of *SpatialThingTypeByShape* (KnowRob)
    - may have spatial properties with objects such as being near of some object etc. (KnowRob)
    - has affordance of drop, grasp, move, push, shake, throw (Concept Web)
    - has probability of some adjective properties: hard, silent, tall, thin, round (Concept Web)

- **grasp**
    - is subconcept of *HoldingWithOneHand* (KnowRob)
    - has subconcept *PowerGrasp, IntermediateGrasp, PrecisionGrasp* (KnowRob)
    - is grounded on iCUB robot (Concept Web)

- **push**
    - is subconcept of *TransportationEvent* (KnowRob)
    - might be directional such as push left/right/forward/backward (Concept Web)
    - is grounded on iCUB robot (Concept Web)

3 Conclusion and Future Work

In this report, we propose to combine several existing knowledge-bases together into one knowledge-base, in order to bring together their complementing aspects.

As future work, we plan to demonstrate the usefulness of such a knowledge-base in many robotic scenarios. One important application of this knowledge-base is going to be to define a context model on top of it.

Acknowledgments

For the experiments, we acknowledge the use of the facilities provided by the the Modeling and Simulation Center of METU (MODSIMMER). This work is funded by the Scientific and Technological Research Council of Turkey (TÜBİTAK) through project no 215E133.

References

[1] Celikkanat, H., Orhan, G., and Kalkan, S. (2015). A probabilistic concept web on a humanoid robot. *IEEE Transactions on Autonomous Mental Development*, 7(2):92–106.

[2] Saxena, A., Jain, A., Sener, O., Jami, A., Misra, D. K., and Koppula, H. S. (2014). Robobrain: Large-scale knowledge engine for robots. *arXiv preprint arXiv:1412.0691*.

[3] Singh, P., Lin, T., Mueller, E. T., Lim, G., Perkins, T., and Zhu, W. L. (2002). Open mind common sense: Knowledge acquisition from the general public. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pages 1223–1237. Springer.

[4] Speer, R. and Havasi, C. (2013). Conceptnet 5: A large semantic network for relational knowledge. In *The People's Web Meets NLP*, pages 161–176. Springer.

[5] Tenorth, M. and Beetz, M. (2009). Knowrob-knowledge processing for autonomous personal robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4261–4266. IEEE.

[6] Yao, B., Jiang, X., Khosla, A., Lin, A. L., Guibas, L., and Fei-Fei, L. (2011). Human action recognition by learning bases of action attributes and parts. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1331–1338. IEEE.

[7] Zhu, Y., Fathi, A., and Fei-Fei, L. (2014). Reasoning about object affordances in a knowledge base representation. In *European conference on computer vision*, pages 408–424. Springer.