Institute for Cognitive Systems
Department of Electrical and Computer Engineering
Technical University of Munich

Alle@Home team -Open Platform League, RoboCup@Home 2017

Karinne Ramirez-Amaro, Ismail Lahkim Bennani, Patrick Grzywok, Emilia Lozinska, Yumin Sun, German Diez Valencia, Matthias Humt, Ilya Dianov, Rogelio Guadarrama, Pablo Lanillos, Emmanuel Dean-Leon, and Gordon Cheng

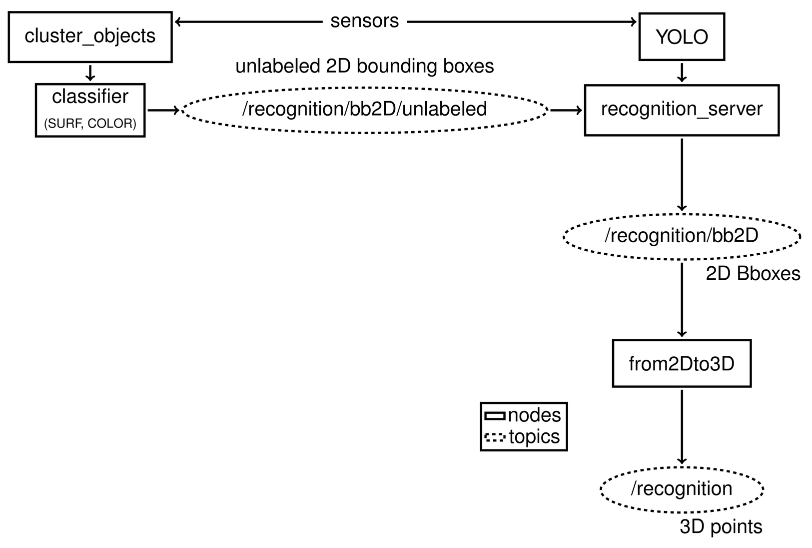
Abstract: We developed a system that integrates different capabilities such as reasoning & knowledge representations, object recognition, navigation, kinematic control, speech recognition and face detection. These capabilities allow the service robot TIAGo to perceive and recognize objects from different places, also to learn, recognize, and follow a new operator. The presented system is fully enhanced with our introduced reasoning and knowledge module to infer incomplete information from observations or verbal commands. We also present a novel control framework that considers information from our developed robot skin to improve the manipulation skills of our robot. Furthermore, we present a novel hierarchical approach that is capable of inferring on-line newly demonstrated everyday human activities from virtual scenarios.

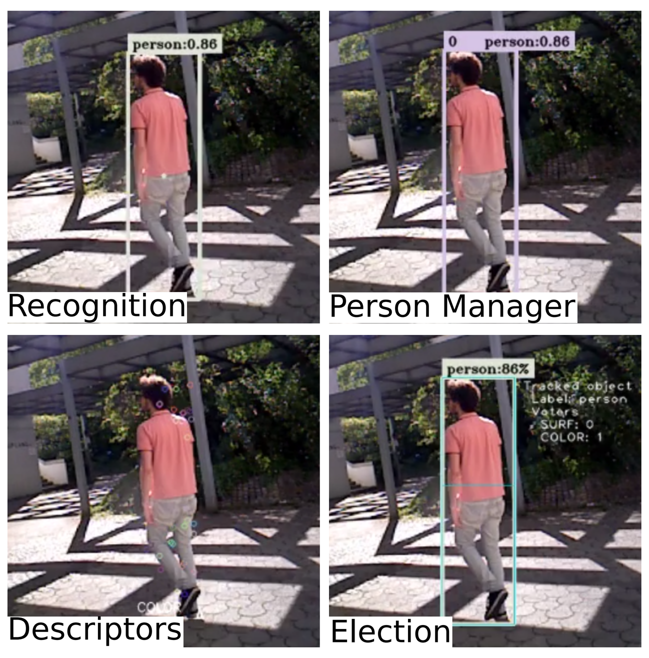
Perception Module

Goal: Observe and **recognize objects** using RGB-D input to compute their 3D position in a real environment.

- Detect objects using YOLO [2] a **deep convolutional network**, which is **robust** recognizing multiple objects.

- Enhance the recognition of objects applying an **Euclidean clustering** method based on depth data.





Goal: Learn, identify & **track the operator**.

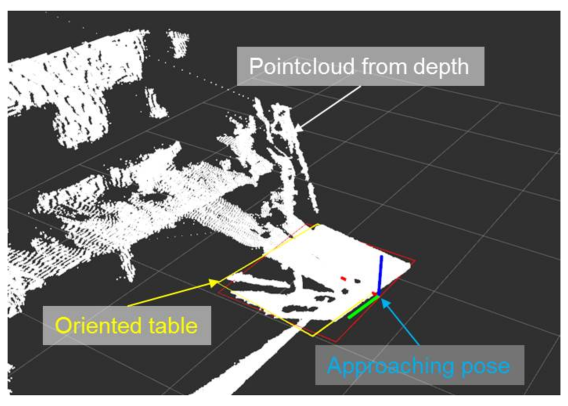
- To update the environment of people, we treat that as a **stable marriage problem** using the Gale-Shapley algorithm [3].

- To detect the operator, we use RGB images trained with **hue histograms** and **SURF** descriptors.

Goal: Identify **tables** in the environment.

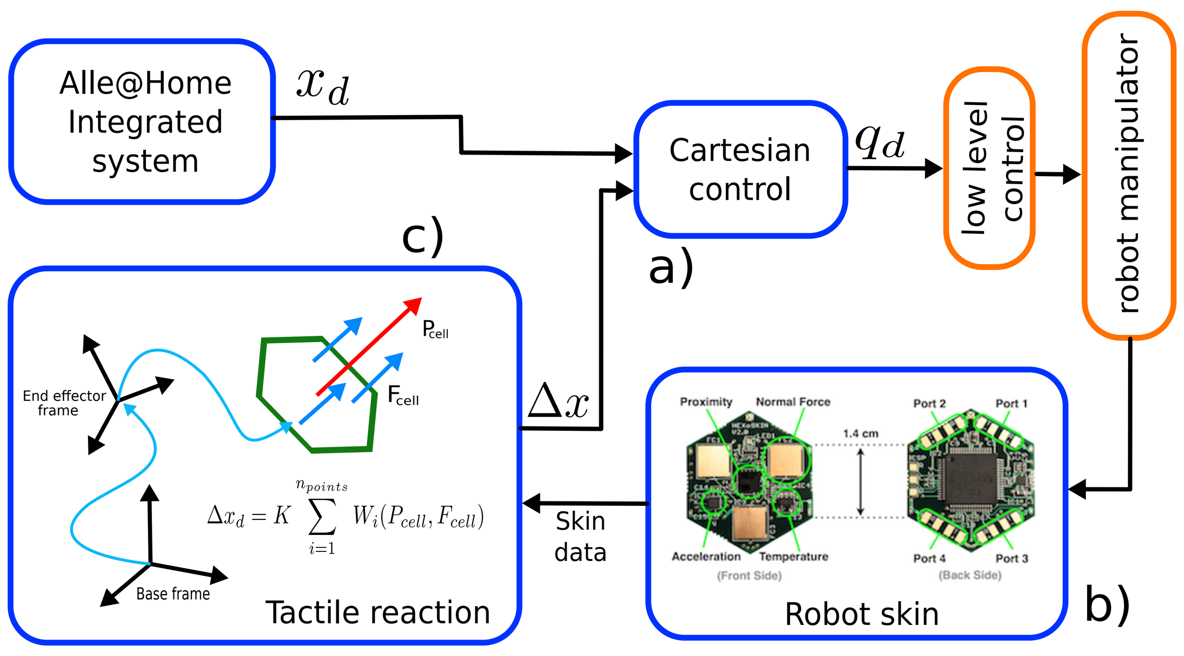
- The RGB-D image is used to **segment planes** from the scene using RANSAC.

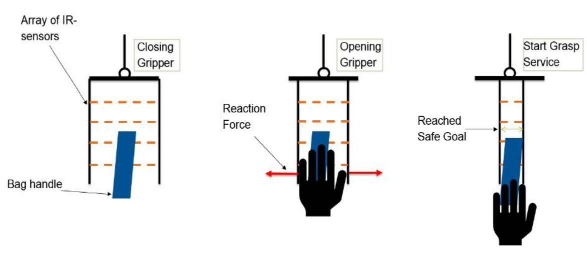
- A **convex-hull** is computed to obtain the boundaries of the plane.



Control Module

Goal: Design **controllers to manipulate** different objects from different locations, e.g. **grasping** from a table and **placing** on a shelf. We develop a library of low-level controllers to produce different robot behaviors applicable to different robots [5].



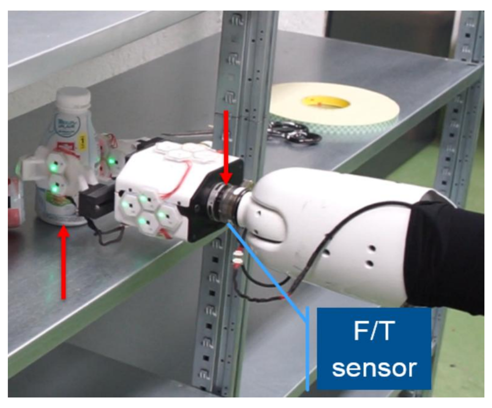


Interaction of the human hand and the gripper during bag handover.

a) A cartesian controller tracks the desired positions and orientations commanded by the higher level nodes.

b) **Robot skin** is used to sense distance from objects (**P**i) and contact forces (**F**i) in the end effector in a n-cells network.

c) Skin signals Fi and Pi of Celli are used to compute the interaction forces, which are used to shift the desired cartesian position to make grasping and placing safer.



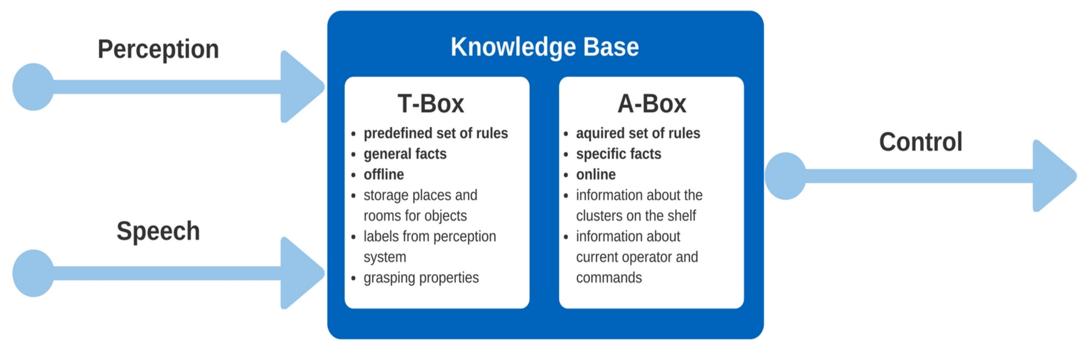
Examples of external forces applied to the system between the gripper, the object and the shelf.

Reasoning and Knowledge Module

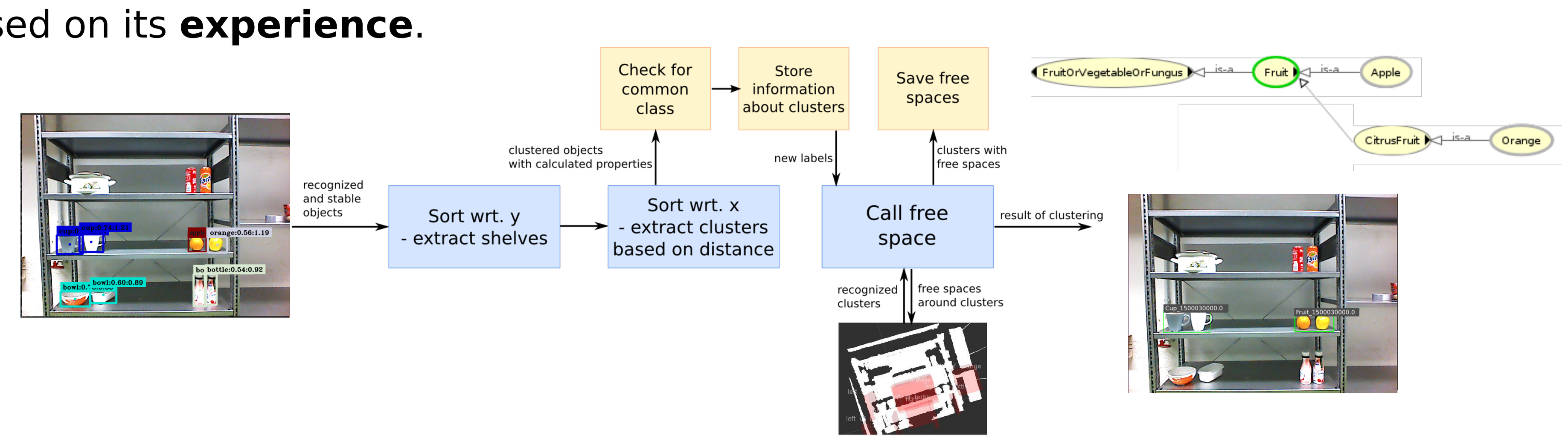
Goal: Infer incomplete information from several sources of information

- Detect **relationships** between objects using general facts (**TBox**).

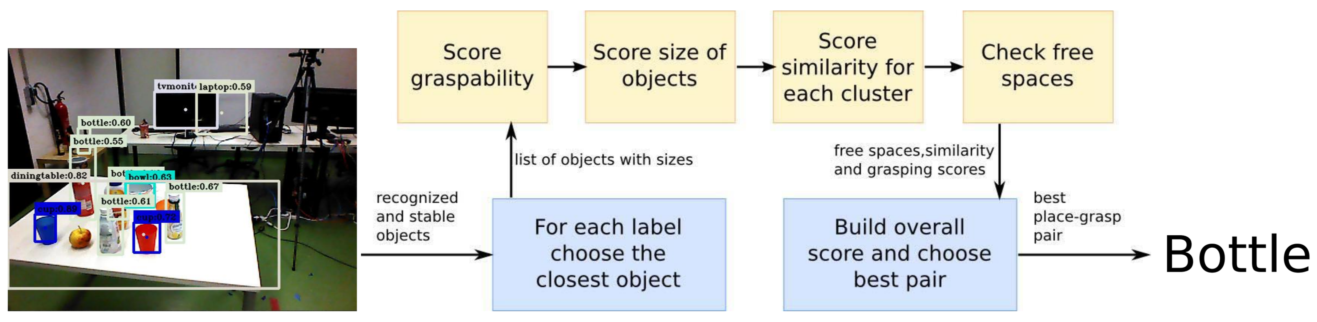
- Properly store the acquired knowledge from tasks (**ABox**) to **make informative decisions**.



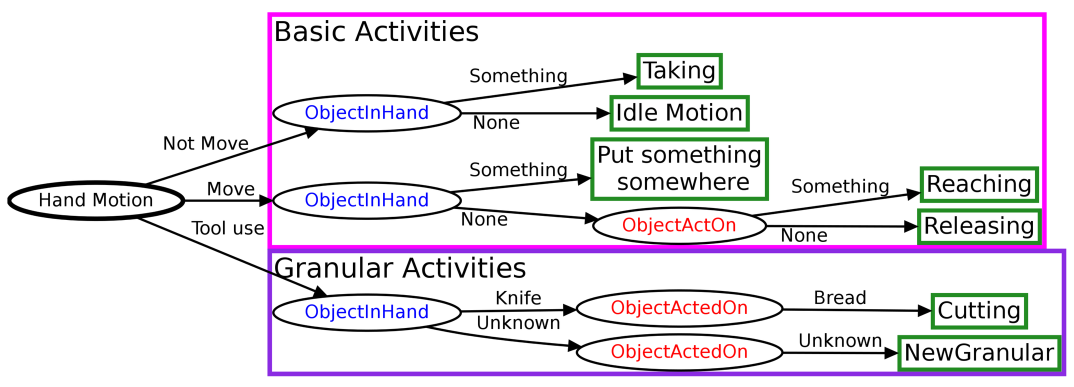
For example, we are able to **cluster** the recognized objects based on their **properties**. Then, the system makes a decision about the best graspable object based on its **experience**.

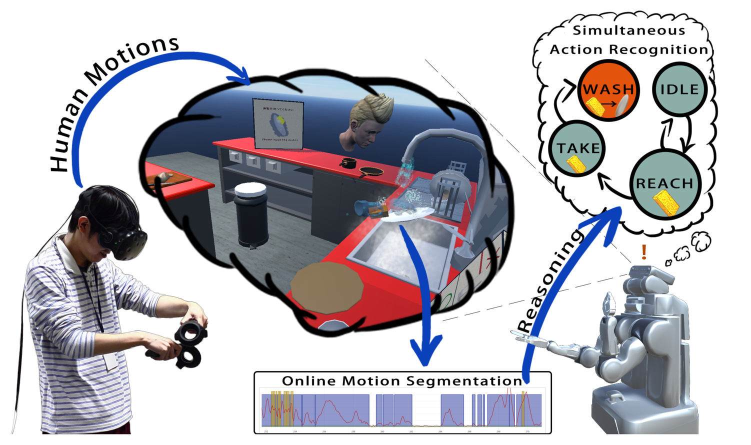


Retrieving information from the knowledge base to **infer the best object to grasp**.



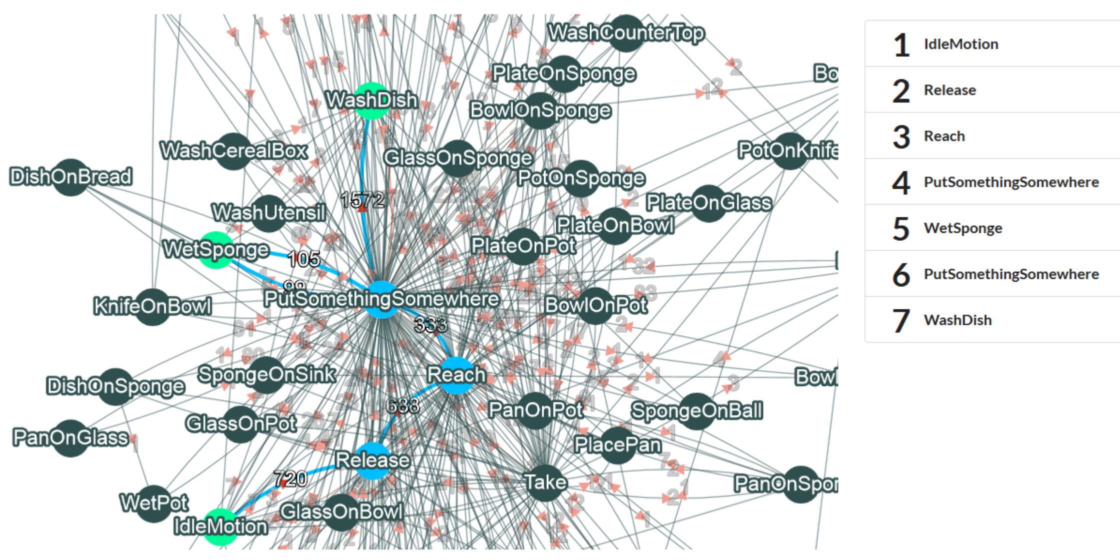
We present a **hierarchical approach** to extract the meaning of demonstrations by means of **symbolic and semantic representations** [1]. These **general common representations** are used to generate a semantic reasoning engine to **transfer** the obtained models among different domains [6]. Our reasoning-based learning system allows robots to **re-use their previous experiences** to correctly segment and **recognize newly** demonstrated activities for new tasks [1,4].





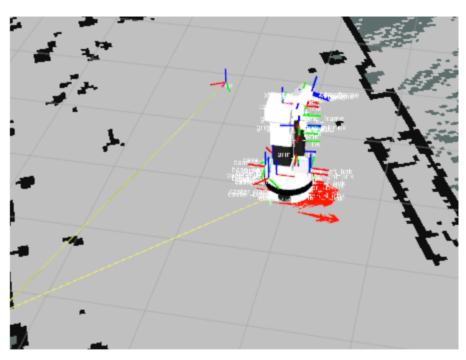
- A new method to teach robots **new tasks** in a fast and efficient manner by **extracting key structures** from the shown tasks [4].

- Automatically obtain the task relationships as a directed graph to **capture** important task components.



Navigation Module

We use the ROS navigation stack enhanced by the reasoning system. For building the map we use Karto and for the localization, we use the **Adaptive Monte Carlo Localization** approach with laser scanner data. We introduced 3 obstacle layers for sonars, lasers, and RGBD camera.



References

[1] Karinne Ramirez-Amaro, Michael Beetz, and Gordon Cheng. "Transferring skills to humanoid robots by extracting semantic representations from observations of human activities". Artificial Intelligence, Special Issue on {AI} and Robotics, Vol. 247, Pages: 95-118, 2017.
[2] Joseph Redmon and Ali Farhadi. "Yolo9000: Better, faster, stronger". arXiv preprint arXiv:1612.08242, 2016.
[3] D. Gale and L. S. Shapley. College admissions and the stability of marriage. The American Mathematical Monthly, 69(1):9–15, 1962.
[4] Tamas Bates, Karinne Ramirez-Amaro, Tetsunari Inamura, Gordon Cheng. "On-Line Simultaneous Learning and Recognition of Everyday Activities from Virtual Reality Performances". IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017.
[5] Emmanuel Dean-Leon, Florian Bergner, Karinne Ramirez-Amaro and Gordon Cheng. "From Multi-modal Tactile Signals to a Compliant Control". IEEE-RAS International Conference on Humanoid Robots, 2016.
[6] Karinne Ramirez-Amaro, Emmanuel Dean-Leon, Ilya Dianov, Florian Bergner and Gordon Cheng. "General Recognition Models Capable of Integrating Multiple Sensors for Different Domains". IEEE-RAS International Conference on