Towards Robust Human Robot Collaboration in Industrial Environments

Batu Akan, Baran Çürüklü, Giacomo Spampinato, Lars Asplund
Mälardalens Högskola, Box 883, 721 23 Västeras, Sweden
{batu.akan, baran.curuklu, giacomo.spampinato, lars.asplund}@mdh.se

***Abstract*—In this paper a system, which is driven through natural language, that allows operators to select and manipulate objects in the environment using an industrial robot is proposed. In order to hide the complexities of robot programming we propose a natural language where the user can control and jog the robot based on reference objects in the scene. We used semantic networks to relate different types of objects in the scene.**

***Index Terms*—Human robot interaction, Object selection, Robot collaboration**

I. Introduction

In order to make industrial robots more favorable in the small and medium-sized enterprise (SME) sector, the issues of flexibility has to be resolved. Typically for those SMEs, that have low volume production and frequently changing applications, it is quite expensive to afford a professional programmer or technician, therefore a human robot interaction solution is strongly demanded in order to let the user to program these systems in an intuitive way. Using a high-level natural language, which hides the low-level programming from the user, will enable a task expert who has knowledge in manufacturing process to easily program the robot and let the robot to switch between previously learned tasks. Thus, the goal is to eventually bring robot programming to a stage where it is as easy as teaching the task to a new member of the work team.

Addressing the objects in the environment by means of natural language can be problematic. If every object in the working environment or on the workbench has different properties such as color or size then these objects can be addressed using these properties as references, but in a typical industrial environment many of the objects may look similar or have very similar properties but have different functionalities. One possible solution is to use spatial terms such as *in front of* or *next to* as a reference between the objects [5], [2], However this approach maybe confusing and time consuming in cluttered workspaces, for the operator. Therefore an hybrid approach that would allow the robot give pseudo-names to the remainder set of objects can speed up the object selection phase and can clear up any confusions, which is vital for a well functioning human robot collaboration.

In this paper a system is introduced which makes use of semantic networks to assign runtime pseudo names to the objects in the scene based on their functional properties per request by the user.

This project is funded by Robotdalen, VINNOVA, Sparbanksstiftelsen Nya, EU European Regional Development Fund.

II. Architecture

The proposed system has there main parts; Automatic speech recognition, visual simulation environment, knowledge base and reasoning system.

In this project Microsoft Speech API 5.1 (SAPI 5.1) [4] has been used, both to recognize the commands given to the robot and to synthesize speech in order to give feedback to the user. The commands are parsed and passed to the reasoning system which eventually drives the physical robot to perform the necessary actions and commands requested by the user. The reasoning system takes the semantic output of the natural language processing unit and tries to find out a possible solution from the set of reasonable actions. Through sequencing low-level skill functions such as *approach*, *pickup*, *release*, reasoning system creates the robot program to be executed. The simulation environment which can calculate forward and inverse kinematics of a general 6 Degree of Freedom (6DoF) industrial robot is used to verify the reachability of the target objects and collision between objects. It also acts as a visual feedback to user displaying the intentions of the robot. Such as which object is going to be picked up or which path will be taken. If the operator does not like the generated output or if the speech recognition system has miss heard the user then the user can revoke previous commands from the execution queue [2].

A. Semantic Networks

A semantic network is a network which represents semantic relations among concepts [3]. With in the artificial intelligence (AI) community it is often used as a form of knowledge representation and is commonly useful for natural language processing. As seen in Figure 1 all objects in the system are derived from a base type named *Object*. Within this framework the objects around the robot are put into five sub categories; (i) manufacturing parts, (ii) movable tools, (iii) non-movable machines used by robots and/or humans, (iv) working benches, (v) miscellaneous objects [1]. Upon recognition, each object in the scene is assigned to one of the relevant categories or sub-categories derived from these main categories, while inheriting the default properties of this category. Additional properties can also be defined that is only relevant to this object or base properties can be overridden.

Assigning the objects to a functional group is very useful in several ways. It limits the operations that are allowable on the objects, therefore eliminating the risks of the robot to perform irrelevant tasks, such as; picking up a workbench or trying to

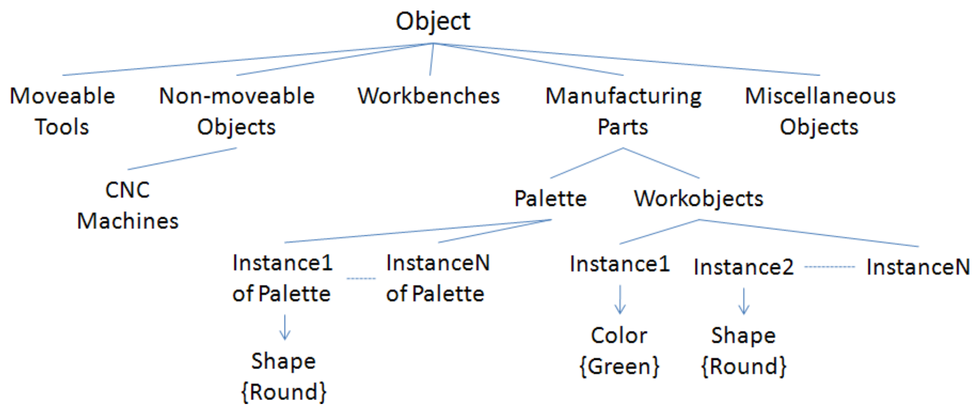

Fig. 1. Representation of the semantic network. All categories are derived from *object* and can have properties different meta properties.

palletizing a CNC machine. It increases the overall robustness and security of the system.

Similarly, category and sub-category names can be used to assign pseudo-names to the objects in the scene. Upon saying *"Please name <object_category>"* Reasoning system scans through all the objects in the memory and assigns a pseudo-name to each; beginning with the category name appended by a greek letter.

III. Experimental Results

In order to demonstrate the usefulness of the proposed system. An ABB IRB-1600 robot which has been equipped with a camera system has been used. A calibrated workbench is used. When the camera system scans over the workbench and recognizes the objects type and returns their positions and orientations in workbench coordinate system. Each recognized object is added into the knowledge base and the associated CAD model is shown on the visual simulation environment. A sample dialog between the robot and the operator is given below and Figure 2 presents a screenshot of the simulation environment.

User:	Can you scan the workbench for objects?
User:	Execute
Robot:	Done.
User:	Please name the palettes.
Robot:	The name of this object is palette beta,
Robot:	The name of this object is palette gamma
User:	Can you put all blue objects into palette beta?
Robot:	*(Checks for reachability)* Yes, I can.
User:	Execute
Robot:	*(picks up the blue objects and puts them into palette gamma)*
User:	Show me palette beta.
Robot:	*(Goes over palette beta)*
:	It is this one.
User:	Can you put all red objects into palette beta?
Robot:	Yes, I can.
User:	Execute
Robot:	*(picks up the red objects and puts them into palette beta)*

IV. Conclusion

In this paper, we demonstrated a high-level restricted language in order to command an industrial robot for simple pick and place applications. The proposed language can handle attributes of the objects in the environment, such as shapes, colors, and other features using a semantic network. It can also use this knowledge base to assign pseudo-names to the objects in the scene to allow fast object referencing.

References

[1] B. Akan, B. Curuklu, and L. Asplund. Interacting with industrial robots through a multi-modal language and sensor systems. Seoul, Korea, October 2008. International Symposium on Robotics.

[2] B. Akan, B. Curuklu, G. Spampinato, and L. Asplund. Object selection using a spatial language for flexible assembly. Mallorca, Spain, September 2009. Emerging Technologies on Factory Automation, ETFA09.

[3] I. Bratko. *PROLOG Programming for Artificial Intelligence, Second Edition*. Addison-Wesley, 1990.

[4] http://www.microsoft.com/speech. *Microsoft Speech Technologies*.

[5] B. Landau and R. Jackendoff. What and where in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16:217–265, 1993.

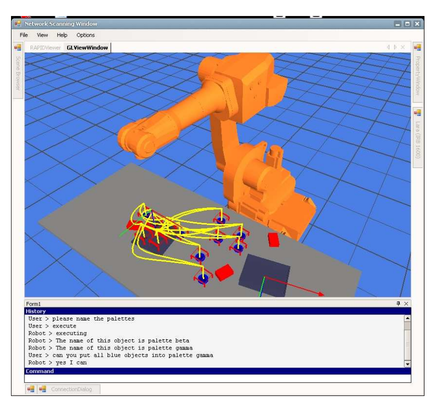

Fig. 2. Screenshot of the simulator where the robot is asked to pick and place all the blue objects into palette gamma