# **Multi-aspect natural language interface to robots:** the R5-COP iTrolley warehouse robot demonstration

Tamás Mészáros, Péter Györke, István Engedy, Péter Eredics, Tadeusz Dobrowiecki Budapest University of Technology and Economics, Budapest, Hungary

### Summary

We present an experimental user interface based on natural language voice understanding and generation enhancing human-machine communication in robotic systems. The system was implemented in the framework of the R5-COP (Reconfigurable ROS-based Resilient Reasoning Robotic Cooperating Systems) project.

The interface is based on a limited set of natural language grammatical rules and vocabulary that can be dynamically composed and reconfigured according to the tasks and skills of the robotic system. The language covers different aspects of the robotic system including sensors, actuators, motion, goals, actions etc. for which language components can be developed independently in a modular way. The final interface is composed dynamically according to the actual configuration and operation of the robot.

The system was developed and integrated using the Robot Operating System (ROS). A communication component called Voice Agent provides a bidirectional natural language voice communication pertinent to the user, the robot and the task at hand. This agent routes input and output between the human user and the robot's internal components. It translates voice input into ROS messages and publishes them to their appropriate ROS topics. Other agents listen to these topics and perform tasks using robot's sensors and actuators. They can also engage in natural language conversations with the human user to clarify the situation.





## Architecture and system components

The entire application is designed as a ROS-based system. The ROS platform provides the basic organization and communication capabilities that connect the robotic components as well as other application specific modules.

In order to design the overall architecture we have selected the intelligent agent methodology. This issue has already been investigated in the R5-COP project and various individual characteristics of the agents (especially autonomy and reactive behavior), their communication

<pre>type : subscripe, : sender : "bickupAgent", tecgrition_topic : "PickupAgent_SpeechRecognition", accepted_patterns : { regexp : ".*what are you doing.*", priority : "1" } } </pre>	text : "Please follow me to pick up the 'beck chair' called tilifaile." } 11:56:36.012 { type : "request", sender : "PlckupAgent", target : "Please locate the box containing the deck chair named tillfalle, and r }	target: "SnoppingListAgent_Control", "product": [filler", "type": "RECOP-9-", "type": "deck chair", "price": 1220.0", "fos_X": "-400", "pos_X": "-400", "pos_X": "S80" }, "status": "at_product" }
ItemCollectorAgent_SpeechRecognition	ItemCollectorAgent_Additem	ItemCollectorAgent_PickupSuccess
target : "ItemCollectorAgent_SpeechRecognition", content : "start collecting items" }	target : "ItemCollectorAgent_Additem", content : "RSCOP-8" }	sender : "PickupAgent", target : "PickupSuccess", "product" : { "name", "blaungen", with
11:55:27.682	····11:54:51.437 ····	"type": "clothes peg", "price": "5.99", "discount": "15"
type : "inform",	type : "inform",	"pos_x" : "-400",
sender: "SpeechkecognitionNode", target: "ItemcollectorAgent_SpeechRecognition", content: "what are you doing" }	sender: "SnoppingListWode", target: "ItemCollectorAgent_Additem", content: "R5COP-9" }	pos_yr : "400" }. "status" : "pickup_success" }
PickupAgent_Arrival sender : "ItemCollectorAgent".	PickupAgent_Pickup	R5COP_Management
target : "PickupAgent_Arrival", "product" { "name": "tillfalle", "d" : "RSOP-9", "type" : "deck chair", "price": 129.0", "discount": 10", "pos_x": -400", "down",	<pre>11:56:14.951 { type : "inform", sender: "ShoppingListNode", target : "PickupAgent_Pickup", content : "RSCOP-8"</pre>	
<pre>pus_y: 300 }. "status": "arrival_at" }</pre>	\$	

# Pilot Use Case: the iTrolley system

To demonstrate the design and operation of the natural language interface a sample scenario of a prototype (virtual) robot called iTrolley working in a self-service warehouse environment was developed. Its main task is to help customers to pick up and carry (large) boxes of the selected wares. A typical application is a furniture store where the user selects goods in the display area and picks up the wares in a self-service storage area.

After assembling a shopping list the robot helps the user to traverse the storage area and to pick up the selected goods. During this it plans for the optimal route, instructs the customer about what boxes should be picked up, resolves problems and warns about mistakes. It may also provide additional services like payment preparation, customer support related to the wares, marketing, etc.

The pilot system is implemented using the Gazebo simulator, Rviz visualization environment and the Jackal robot in a custom-made small warehouse environment. The navigation is done by Odom\_navigation of the Jackal robot. The voice interface is implemented using Google Voice and Speech Synthesis APIs in the Android environment. The Voice Agent runs on Android while other agents and ROS components reside on Ubuntu Linux.



and cooperation capabilities were found to fit the needs of R5-COP robotic systems.

The system internal architecture contains the following main components:

- · Voice Agent running on a mobile device that performs the natural language communication with the user and connects to the ROS-based system.
- · Robotic Modules running on the robot that provide services like physical movement of the robot, handling sensor input and other.
- Local Application Agents running on the robot that provide non-robot related services like natural language processing and understanding, state management of human-robot conversations and other.
- Global Application Agents running on servers and provide services like synchronizing the operation of multiple robots, resolving conflicts, logging and monitoring, etc.



# A shopping scenario using the iTrolley robot

- 1. The Customer enters the Warehouse then
- assembles the shopping list by scanning QR-codes in the display area, or
- asks the robot "what is on sale" and decides to add these items to the shopping list.
- 2. After the shopping list is ready the customer enters the storage area and
- instructs the robot "start collecting items" in order to collect the wares.
- 3. The robot guides the customer in the storage area. For all wares it
- selects the shelf containing the ware closest to the customer,
- plans the route to the shelf and travels to the right place,
- instructs the customer to put the box on the robot transport platform,

# Communication and cooperation

There are three layers of communication with different purposes and requirements:

- ROS messages: low-level communication between all system components
- Agent messages: higher-level communication between autonomous entities
- Natural language messages: high-level communication with the human user

At ROS-level the message content and the cooperation is determined by the requirements of the applied ROS modules.

At the agent-level messages follow the Agent Communication Language (ACL) standard that uses performatives to express the nature of the communication and content models to transfer the information between communicating parties. We developed several cooperation protocols to cover various aspects of the prototype application including shopping list assembly and ware pickup.

At the highest level the content of ACL messages is described using task-specific controlled natural languages. These are easy to use by humans and they are also easy to process and understand by agents.

- checks the QR-code on the box and warns the customer if it is not correct.

4. After all items from the shopping list are collected, the robot - notifies the customer that all boxes have been found, and

- moves to the checkout area to proceed with the payment.

# **Evaluation and further development**

The developed system provides a natural language voice interface to a robotic system that provides a complex functionality to the human user. The interface language is composed from a set of task-specific controlled natural languages thus creating a multi-perspective communication that suits the robot's actual configuration and tasks. The system is also able to engage in conversations with the user even on different topics at the same time.

Although the system was developed for the iTrolley pilot scenario it can be easily adapted to other applications as well by developing suitable application-specific agents and task-specific natural languages. The current demo is voice-based but other modalities could also be used by developing other interface agents.

# Contact

Tamás Mészáros <meszaros@mit.bme.hu> http://r5cop.mit.bme.hu/

