Natural Language Processing in Domestic Service Robotics

Stefan Schiffer

Knowledge-based Systems Group
RWTH Aachen University

poster at

LingUnite – Tag der Sprachforschung

11. Oktober 2013, RWTH Aachen University

Abstract

As robots are more and more entering our everyday life, like as assistive devices to support us in our home with the daily chores, methods to control and to interact with such robots become more and more important. The most natural way for a human to instruct a robot is perhaps natural language. However, there are several challenges to master to allow for suitable human-robot interaction by means of natural language. We will report on two of our efforts in enabling humans to use natural language to command a domestic service robot. The two methods we present reside on different levels - one is at a lower level of recognizing speech from acoustic input while the second one is about interpreting natural language. While the former was primarily intended for noisy scenarios to help rejecting utterances that were not meant for the robot, the latter yields a flexible system for commanding a robot which can resolve ambiguities and which is also capable of initiating steps to achieve clarification.

The first approach [1] is at the signal processing stage where the acoustic input received from spoken language is to be converted to the textual level. When acting in human environments it is important that commands given to the robot are recognized robustly. Also, spoken language not directed to the robot must not be matched to an instruction for the robot to execute. We developed a system that is robust in noisy environments and that is

insusceptible to act upon commands not meant for the robot. First, we use a threshold-based close speech detection to segment utterances targeted at the robot from the continuous audio stream recorded by a microphone. Then, we decode these utterances with two different decoders in parallel, namely one very restrictive decoder based on finite state grammars and a second more lenient decoder using N-grams. We do this to filter out false positive recognitions by comparing the output of the two decoders and rejecting the input if it was not recognized by both decoders.

The second approach [2] takes place on a higher level of abstraction, that is, it deals with interpreting an utterance that has already been transformed to text from the raw audio signal. We model the processing of natural spoken language input as an interpretation process where the utterance needs to be mapped to a robot's capabilities. More precisely, we first analyse the given utterance syntactically by using a generic grammar that we developed for english directives. Then, we cast the interpretation as a planning problem where the individual actions available to the planner are to interpret syntactical elements of the utterance. If, in the course of interpreting, ambiguities are detected, the system uses decision-theory to weigh different alternatives. The system is also able to initiate clarification to resolve ambiguities and to handle errors as to arrive at a successful command interpretation eventually. We show how we evaluated several versions of the system with multiple utterances of different complexity as well as with incomplete and erroneous requests.

References

[1] Masrur Doostdar, Stefan Schiffer, and Gerhard Lakemeyer. Robust speech recognition for service robotics applications. In *Proceedings of the International RoboCup Symposium 2008 (RoboCup 2008)*, volume 5399 of *LNCS*, pages 1–12. Springer, July 14-18 2008. Best Student Paper Award.

[2] Stefan Schiffer, Niklas Hoppe, and Gerhard Lakemeyer. Natural language interpretation for an interactive service robot in domestic domains. In Joaquim Filipe and Ana Fred, editors, *Agents and Artificial Intelligence*, volume 358 of *Communications in Computer and Information Science*, pages 39–53. Springer Berlin Heidelberg, 2013.