Cooking with a robot: The Dialogue System

Intelligent Systems Laboratory — Winter Term 2013/2014

Alexander Vorwerg & Monika Chromik & Patryk Pekala

Supervisors: Anja Philippsen & Birte Carlmeyer & Andreas Kipp & Britta Wrede

Bielefeld University, Faculty of Technology

Universität Bielefeld

Abstract

This project is part of a kitchen assistance program providing different user interfaces. The role of this project is the voice interface. The user should be able to get kitchen device statuses, cooking instructions and explanations about the use of the kitchen and cooking procedures using this dialogue interface. In this way the user has hands and eyes free for the cooking process and does not have to worry about timing or statuses of different kitchen devices.

Introduction

In a modern kitchen, many different devices are connected to one operation system which provides an overview of all accessible functions. Additionally the devices help the user to interact with them and to react precisely if a specific action has to be performed. During a regular cooking process the user has to coordinate several tasks, e.g. preparing ingredients, timing processes and monitoring temperature. Furthermore the user has to study the recipe as well which can be intimidating for novice cooks. So a stable and easy to access way of getting this information and interacting with the kitchen is needed. The cooking process itself occupies both hands and the eyes, but the mouth is free to talk and the ears are not occupied as well. So the communication via speech input and output seams to be a good choice. Furthermore it is a familiar way of getting assistance, since human assistance works the same way if not physical. Speech has also the convenience that it is accessible from all over the kitchen without any further effort. But the voice output includes the difficulty that only one information can be given at a time, so the task of selecting and portioning the information can not be done by the user and has to be provided by the system.

This project is meant to be integrated in a system with other input and output resources to extend the user experience and evaluate different access choices.

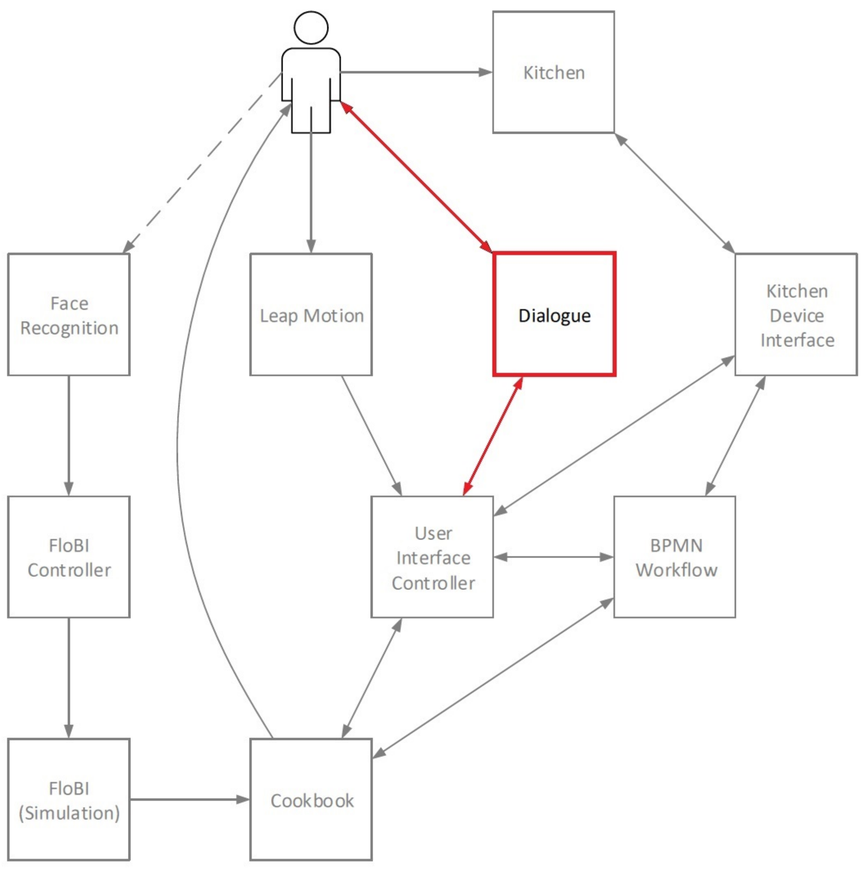

Figure 1: Integration in the CWAR(Cooking with a Robot) environment. The dialogue is an interface between the user and the general user interface controller. The interface controller is the access to the cookbook and kitchen sensors.

Main Objectives

- Dialogue system to interact with kitchen
- Step by step cooking assistance
- Explanation of kitchen functions
- Communication of monitored events
- Stable speech recognition
- Introduce novice users to cooking
- Enhance cooking experience
- Avoid kitchen misuse

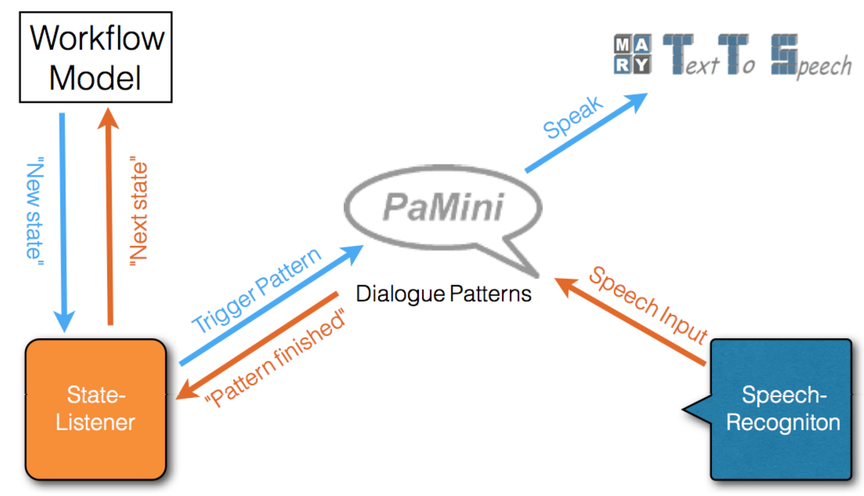

Figure 2: Internal structure of the dialogue system. The heart piece are the PaMini dialogue patterns, handling the ISR speech recognition and the changes in the internal states caught by the state listener. The state listener then triggers new dialogue patterns and requests information by the operating system, corresponding to the finished patterns.

System Architecture

The center piece for our project is the PaMini framework (Pattern-Based Mixed-Initiative Human-Robot-Interaction) [1]. The provided concept of generic interaction patterns *(Figure 3)* allows the system to react fluently to questions of the user or sudden events triggered by sensors of the kitchen and then to resume in the current main path of assistance *(Figure 2)*. For the input we use the ISR speech recognizer (based on Esmeralda [4]) which streams the results in real time to our dialogue manager via RSB (robotics service bus) [2]. This input is interpreted by the dialogue manager which triggers corresponding patterns or resumes in the activated ones. Input from the user is then reported to the ”cookbook” to get the next cooking instructions or the necessary information about the status of the different kitchen devices. As an output resource of the dialogue patterns we use the text-to-speech system MaryTTS [3].

Although this setup could work as a stand alone communication to the kitchen and its sensors, it is planned to be integrated in a system with alternative input and output resources like Leap Motion for the input and a screen for an output, which are provided by other Cooking-with-a-robot projects.

The platform where this project is meant to be running is a Miele@Home kitchen which provides the necessary information from the different kitchen devices bundled.

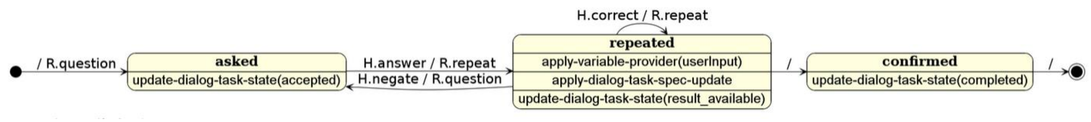

Figure 3: The graph of the PaMini pattern for the robot requesting information and expecting confirmation of the repeated information, or otherwise repeating the question.

Results

At this point we have a stand alone prototype supporting input via speech recognition as well as input via GUI for testing and to bypass possible problems with the speech recognition.

The voice output is supported by text output on the screen, to monitor the output history. The information later provided by the kitchen is currently hard coded in the systems database, since the integration in the kitchen system is yet to come.

The quality of speech recognition needs an improvement, particularly because of the more challenging noise situation during the cooking process.

The system provides a step by step cooking assistance which can handle interrupting user questions concerning the kitchen or cooking techniques without any disturbance in the procedure. But handling of interruptions caused by kitchen events like overheating, etc. has to be added when the system will be integrated in the operating system.

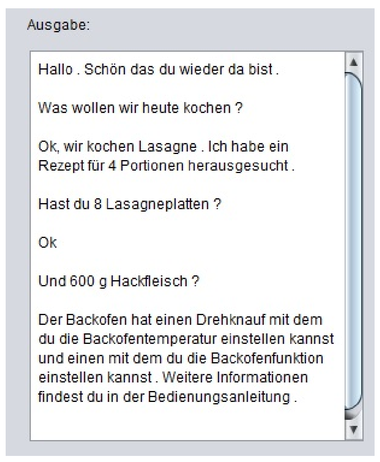

Figure 4: Example for the robot‘s output, showing the beginning of a conversation and the interruption of cooking advice by answering to the user, asking how to use the oven.

Prospects

- Integration of the system in the real kitchen
- Coordination with visual output
- Improvement of speech recognition
- Cooking suggestions
- Proactive proposal of help by the robot
- Entertainment options on waiting time
- Evaluation of dialogue versus Leap Motion & screen

Conclusion

At this point it is plausible that communication with the kitchen and its devices via voice input and output is a good choice, since the hands and eyes are still free for cooking procedures and the system is flexible enough with the dialogue to answer to user questions while giving cooking instructions. The voice recognition should to be improved, since a kitchen in use is a noisy place, but still an assistance system should work stable under that condition, so that it does not require too much attention.

References

[1] Julia Peltason and Britta Wrede. Pamini: A framework for assembling mixed-initiative human-robot interaction from generic interaction patterns. In *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 229–232. Association for Computational Linguistics, 2010.

[2] Johannes Wienke and Sebastian Wrede. A middleware for collaborative research in experimental robotics. In *System Integration (SII), 2011 IEEE/SICE International Symposium on*, pages 1183–1190. IEEE, 2011.

[3] Marc Schröder and Jürgen Trouvain. The german text-to-speech synthesis system mary: A tool for research, development and teaching. *International Journal of Speech Technology*, 6(4):365–377, 2003.

[4] Gernot A Fink. Developing hmm-based recognizers with esmeralda. In *TSD*, volume 99, pages 229–234. Springer, 1999.