Paper: “Personalized short-term multi-modal interaction for social robots assisting users in shopping malls”

Authors: Luca Iocchi, Maria Teresa L´azaro, Laurent Jeanpierre, Abdel-Illah Mouaddib

SAPIENZA
UNIVERSITÀ DI ROMA

Nazgul Mamasheva

Personalized short-term multi-modal interactions with non-expert users (customers of a shopping mall)

Fig. 1: COACHES environment and robots.

Personalized means that the robot should use different forms of interactions to communicate the same concept to different users, in order to increase its social acceptability;

Short-term means that the interactions are short and focused on only one particular communicative objective, avoiding long and complex interactions;

Multi-modality is obtained by using different interaction devices on the robot (although in this study, the focus is only on speech and graphical interfaces).

Personalized short-term multi-modal interactions with non-expert users (customers of a shopping mall)

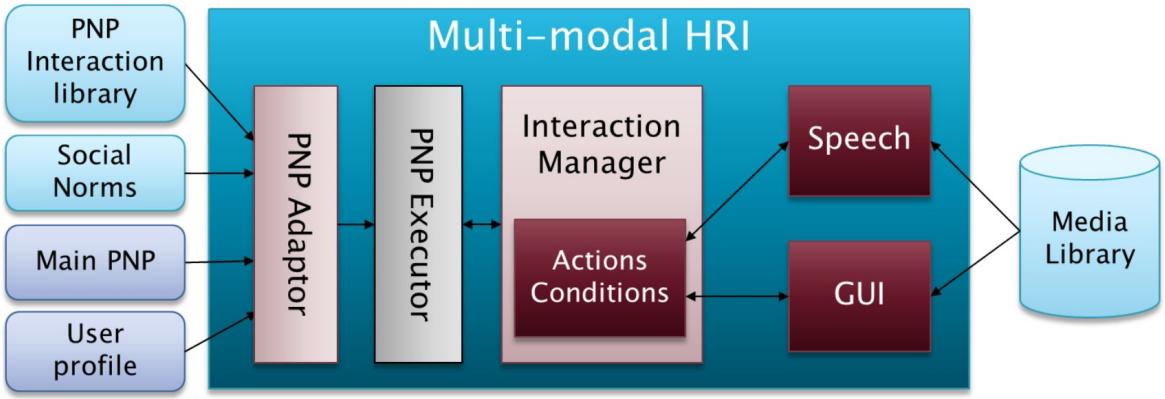

Fig. 3: Architecture of Human-Robot Interaction module.

Main PNP calls Interaction PNPs as sub-routines. All these plans are processed and transformed by applying the social norms (described as rules) customized to the current user profile.

Petri Net Plans (PNPs) encode the overall behavior of the robot (planning and reasoning components of the system), social norms, a user profile and a multi-media library.

Social norms are domain and task independent and are represented using a propositional logic formalism.

(child, use_animation)
(elder, use_big_font)
(elder, use_simple_GUI)
(deaf, ¬ use_speech)
(blind, ¬ use_display)
(elder ∨ deaf, display_spoken_text)
(elder ∨ deaf ∨ blind, ask_for_guidance)
(blind, use_detailed_speech)
(blind, notify_guidance)
(first_time_user, detailed_instructions)
(¬ first_time_user ∧ young, ¬ detailed_instructions)
(child ∨ very_young, ¬ use_baby_care_room)
(foreign, speak_English)

Table 1: Domain-independent social norms.

User profile is the information available about the user that is interacting with the robot. Identification mechanisms: users wearing an RFID tag containing personal information, by providing fidelity card, by entering personal password or by showing QR-code to HRI robot.

PNP Adaptor and Executor

The HRI module is implemented within the framework of the Petri Net Plans (PNP) formalism.

PNPs are based on two main concepts: *actions* (i.e., output operations) and *conditions* (i.e., input operations).

Actions include motion of the robot in the environment, spoken sentences issued by the on-board speakers, text, images, videos or animations shown on the on-board screen, etc.

Conditions include the result of perception routines (e.g., image processing or speech recognition), the input given by a user through a GUI on the on-board screen, information about personal data of user acquired through a reader of fidelity cards, etc.

PNP Adaptor generates a personalized plan, given a main plan, a library of interaction plans, a set of social norms, and the user profile. The generated personalized plan is then executed by the PNP Executor.

PNP Executor treats actions and conditions without giving them any semantics and controls only the flow of execution. The actual execution of the basic actions and conditions is responsibility of the Interaction Manager.

Media Library (ML) is a collection of multi-media data (text, images, animations, video, etc.) that are linked to the communication activities of the robot and to the user profiles.

ML supposed to contain different versions of the same communication target for different users.

For example, icecream advertisement can have a different spoken text and different displayed images or videos for children and adults.

Interaction Manager

The interactions are coordinated by an Interaction Manager (IM), which manages all the robot activities (both the ones related with human-robot interaction and the ones used for implementing the basic robotic functionalities).

For the interaction behavior, actions and conditions are actually related to the Speech and Graphical Interface (GUI) modules.

Speech and Graphical Interfaces

Speech recognition and synthesis
The speech component allows the robot to communicate with humans through vocal interactions. It is formed by Automatic Speech Recognition (ASR) and Text-To-Speech (TTS).

The ASR component analyzes audio data coming from a microphone and extract semantic information about the spoken sentences, according to a predefined grammar. This component allows the robot to understand user spoken commands.

The TTS component transforms text messages in audio data that are then emitted by the speakers on-board the robot. This enables the robot to speak to people.

Graphical User Interface (GUI) component implements a graphical input and output interface between users and robots that is displayed through the touch screen on-board the robot. The GUI defines actions (i.e., output operations) and conditions (i.e., input operations) that are integrated in the IM with other communication primitives (e.g., speech) in order to implement a multi-modal interaction.

	Action	**Condition**
Speech	*Say* speak information though TTS	*ASR* Results of ASR
GUI	*Show* show information on the GUI	*GUI* Results of GUI input

Examples of personalized interactions
The examples are taken from the use cases of the COACHES project.

Example 1. Advertising. One of the tasks of the COACHES robot is to show advertisements to users of the shopping mall.
Personalized behaviors/ interactions:

i) animation instead of videos for children,
ii) big fonts and simple GUI for elderly people, etc.

Example 2. Directions and guiding. The robot is able to give directions and guide people in the mall.
Personalized behaviors/ interactions:

i) for elderly people, a simple GUI shows the direction;
ii) The interaction with a deaf and elder person is made with graphical interface only;
iii) the interaction with a blind person uses only voice.

Example 3. Baby care rooms. Baby care rooms can be used by parents, but must be reserved and they are locked when not in use.
Personalized behaviors/ interactions:

i) new user is fully instructed with detailed instructions about how to use the service;
ii) a user that already used the service a few time ago is given directly the access to the baby care room;
iii) children or very young users will be notified that they are not allowed to use the service.

Conclusions

All these examples are implemented without explicit coding the corresponding behaviors.

The expected **personalized behavior** is the effect of the application of the social norms to the user profile and of the corresponding modifications of the plans that activate actions with proper parameters.

Social norms are not specific of any particular task.
This allows for a high level of extendibility.