µDialBot: MUlti-party perceptually-active situated DIALog for human-roBOT interaction

Fabrice Lefèvre, Timothée Dhaussy, Ahmed Njifenjou, Virgile Sucal, Bassam Jabaian
[1]LIA - Avignon Université, France, prenom.nom@univ-avignon.fr

MOTIVATIONS

Companion robots able to assist people in their everyday life and to communicate with them:

social assistive robots

How to improve acceptance?

Several technological barriers to overcome:

- better ability to move, see and hear in order to naturally communicate with people in various configurations (face-to-face, multi-party, close, distant etc)
- pro-active use of perceptions
- audio-visual strategies to handle interactions

AMBITIONS

New level in exploitation of rich information available with audio and visual data flowing from humans when interacting with robots:

- fusing highly informative **verbal and non-verbal perceptive features** to enhance the robot's decision-making ability such that it can
- switch between **multi-party/group interactions and face-to-face dialogues** where required, and
- take speech turns **more naturally**

Tested and validated with several **use cases in a day-care hospital** unit. Large-scale data collection, complement in-situ tests, to fuel further researches.

Project mainly **funded by ANR and labeled by the Pole SCS**. More information on: https://www.pole-scs.org/projets/mudialbot/

PARTNERS

- **LIA, Coordinator** — Human-machine vocal interactions, decision-making learning
- **Lab Hubert Curien** — Image analysis
- **INRIA Perception** — Audio-visual scene analysis
- **ERM** — Robotic engineering, software integration, data management
- **AP-HP/Hopital Broca** — Healthcare application, psychological survey

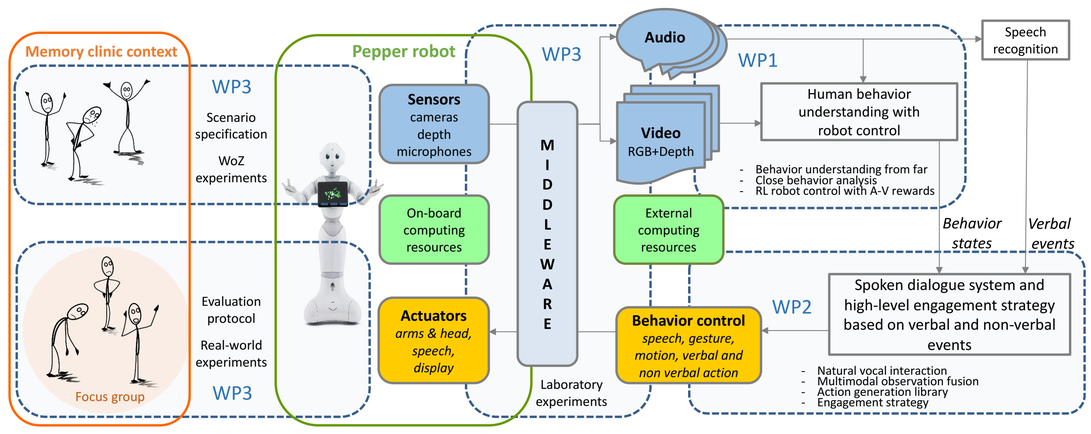
ORGANIZATION

Methodology followed in µDialBot at the same time mainstream, in line with the classical methodology to handle **new machine learning-based approaches**, and pioneering, with **novel online learning techniques** to reduce the requirement for initial data collection to its minimum. Relying on deep learning techniques which have shown their efficiency in other domains, a new formalism dedicated to **pro-active perceptually-based control of a conversational robot** will be proposed.

General methodology of the project consists of two operational building blocks for:

- the estimation of non-verbal states and
- the learning of an event-guided strategy.

Integrated on the robotic platform through a software abstraction layer. A first round of "Wizard of Oz" (WoZ) experiments to test the proposed models, the complete system gradually introduced in the true clinical context using a well-defined protocol.



3 main work-packages implemented:

- ***WP1 Human behavior understanding with robot control***
  Objectives: to develop methods and algorithms to extract HBU cues from audio and visual data.
- ***WP2 Spoken dialogue system and high-level engagement strategy based on verbal and non-verbal events***
  Objectives: to develop the natural vocal interaction ability of the robot.
- ***WP3 Specifications, integration on robotic platform, iterative and final evaluation of the human-robot interactions in a memory clinic***
  Objectives: to define the experimental protocol, specify laboratory and real-world experimental protocols, and conduct "Wizard of Oz" (WoZ) experiments.