numediart
UMONS RESEARCH INSTITUTE FOR CREATIVE TECHNOLOGIES

Cozmo4Resto: A Practical AI Application for Human-Robot Interaction

Kevin El Haddad and Noé Tits

TCTS Lab - numediart institute - University of Mons, Belgium

Objectives

Artificial Intelligence is getting more and more anchored in several aspects of our daily lives such as surveillance systems, medical tools, finance, home/cellphone applications and.... even toys!
In this project we propose to develop interaction interfaces for agents, using machine learning and other approaches, with the goal to apply them in a human-agent interaction experiment. More specifically and as an example, we propose an application based on the robot Cozmo[1]. The goal is to be able to make suggestions of restaurants or meals to a user, based on several factors such as the user's profile, desires, the time of the day and the meals/restaurants available.

[1] https://www.anki.com/en-us/cozmo

Figure 1: (from Anki's website)

numediart
UMONS RESEARCH INSTITUTE FOR CREATIVE TECHNOLOGIES

Work Plan

WP0: Get Cozmo's attention

In order to grasps the agent's attention, systems like keyword spotting, face detection or even a simple voice activity detection will be considered.

WP1: Recognize users based on multimodal cues

Cozmo needs to adapt its responses based on the user. So Cozmo needs to be able to recognize the user. For this we intend to use two main cues, visual and audio. For the visual cue, we plan to use the data of Cozmo's integrated camera for facial recognition. For the audio cue, since Cozmo does, unfortunately, not have an integrated microphone, we plan to use and external one in order to recognize the speaker based on the voice. We also consider to fuse the data in order to give better results.

WP2: Understand the users requests

Not only does Cozmo need to recognize the user (see WP1) but also understand what is being said. For this we will leverage the currently available speech recognition, sentiment analysis and more general NLP systems and adapt them to best suit our application.

WP3: Retrieve relevant data for the task

In order to reply to the user, Cozmo needs to be aware of the context and the available options. Relevant data for this task (such as time of the day, local restaurants, etc.) will be gathered. The tools that will be considered are open source APIs and machine learning-based recommender systems.

WP4: Setting up user profiles

Cozmo should be able to adapt to different users and recognize them after a first interaction. That is why one of the work-packages will be dedicated to build user profiles and update them based on Cozmo's interactions with the users.

OPTIONAL

Some other tasks can be considered and will depend on the amount of participants that will join in this project, as the WPs listed above. These are:

- Recognize and take the users' moods and emotional reactions into account
- Use Cozmo to send messages to people in contact list (invite friends over for diner)
- Take into account restaurants online reviews via sentiment analysis
- etc.

Technical Aspects

A Python SDK is available for Cozmo application development[2]. It allows to control Cozmo's movements and expressions as well as gives access to Cozmo's camera's video data. For the other aspects of this project, APIs (such as Google Maps API) will be leverage to build our application. We will also rely on machine learning libraries such as Tensorflow.

Deliverables & Research Benefits

Our main goal here is to setup a frame work of machine learning-based interfaces for Cozmo applications. Therefore, at the end of this project, we intend to provide the community with a set of machine learning-based tools making the development of Human-Cozmo interactions easier and funnier. We hope this will trigger potential future collaborations and help developing new interesting machine learning-based Cozmo applications.

For Potential Participants

We are looking for motivated participants interested in machine learning, robots and most importantly, **FOOD!!!**. They also should be team players and not be afraid of challenges. So please, if this project motivates you, do not hesitate to apply! More specifically we would prefer participants with **some** of the following skills:

- Python programming, as python will be the main if not only language used in this project (for all WPs)
- Online data harvesting using tools such as Google's APIs or others (mainly for WP3)
- Previous machine learning experience (for all WPs)
- Database management skills in order to store and access data concerning users and relevant for the application (mainly for WP4)
- Dialog management (mainly for WP2 and WP3)
- Natural Language Processing (NLP) (mainly for WP2)

[2] https://developer.anki.com/

numediart
UMONS RESEARCH INSTITUTE FOR CREATIVE TECHNOLOGIES

Profile Team

Kevin El Haddad is a teaching assistant and Ph.D. candidate at the University of Mons. He holds an M.S. in microsystems and embedded systems from the Lebanese University in 2013. His Ph.D. work currently focuses on the use of nonverbal and affective expressions in human-agent interactions. In 2018, he spent 4 months as a lab associate at Disney Research (Los Angeles, California) working on AI and Human-agent Interaction systems. His research interests include machine learning, affective computing, human-agent interactions and signal processing. He lead 3 previous eNTERFACE projects.

Noé Tits obtained his Master of Electrical Engineering specialized in Signals, Systems and Bio-engineering in June 2017. His Master's thesis was done in the context of an Erasmus Plus scholarship at the University of the Basque Country in Bilbao (Spain) in the Aholab laboratory specialized in speech processing. He developed a tool for analyzing pathological voices. His experience also count research projects in the field of electrical engineering such as simulations of heating of cables and electromagnetic fields in cable glands (Laborelec, GDF Suez), motion analysis (eNTERFACE workshop, Numediart Institute of UMONS), singing voice analysis (Hovertone) and Medical Image Processing (in collaboration with UCB). In december 2017, Noé obtained a grant from the FNRS to pursue a doctorate at the Numediart Institute of UMONS. His current research focus on the application of Deep Learning techniques for controlling the emotional expressiveness in Text-to-Speech Synthesis.