SSM
Scuola Superiore Meridionale

MERC PhD Project Proposal 2023/2024

Title of the research project:

Data-driven Architecture for Agent-Human Interactions: From EEGS and LLMs to Models and Decisions

Keywords (up to five)

Autonomous Agents, (Multi-Agent) Reinforcement Learning, Data-Driven Control, Optimal Control

Supervisors (at least two from two different areas):

Supervisors:

Michael Richardson (michael.j.richardson@mq.edu.au)
School of Psychological Sciences, Performance and Expertise Research Centre
Macquarie University, Sydney, Australia
https://researchers.mq.edu.au/en/persons/michael-richardson
Human behaviour, Modelling

Giovanni Russo (giovarusso@unisa.it)
Dept. of Information and Electrical Engineering & Applied Mathematics (DIEM) at University of Salerno
www.sites.google.com/view/giovanni-russo
Reinforcement Learning, Data-Driven Control, Control Theory, Complex Systems, Optimization

Project description (max 5000 characters)

The vision and the context of the project. A pressing challenge is that of designing agents that are able to reliably interact with other agents and humans. To achieve this, it is widely believed that agents must be able to: (i) predict human intentions; and consequently (ii) making decisions based on the states of both the agent and the human. In this context, with this project we will design, test and validate a data-driven architecture for human-agent collaboration. The architecture foresees the use of EEGs to infer the state of the person. This state, eventually processed via LLMs, will be used to infer a causal model of human intentions. In turn, based on this model, data-driven control and learning algorithms will be designed to enable human-agent interactions. For concreteness, we will consider a specific application involving a person interacting with a robotic agent through a collaborative task (e.g., carrying on some equipment) also integrating virtual reality (if needed).

The **results** will be relevant to all applications of complex systems, where agents need to interact across a complex network of connections with humans and other agents.

Project objectives and methodology. The detailed research program will be shaped based on the **interests of students**. A preliminary list of concrete research objectives, with the corresponding methodology, is given below.

O1 –*Literature Review*. This objective is aimed at understanding the basic technologies that will be used and developed within the project. The architecture that we expect to bring to life foresees: (i) EEG headsets to infer the *mental state* of the person; (ii) a signal processing facility that will elaborate the signals from the EEG in a way that these can be parsed by a large language model (LLM); (iii) in turn, the LLM will elaborate these signals breaking them down in elementary instructions that will be used by a robotic agent (controlled via e.g., a data-driven control algorithm); (iv) the robot, upon executing actions, will also interact with the person via LLMs. As such, the very first objective of the project will involve carrying-on a thorough literature review to understand the state-of-the-art of all the technologies involved, highlighting the challenges that the technologies impose on the control, signal-processing components of the architecture.

O2 –*LLM-Mediate-EEG for Control*. This objective has the twofold purpose of: (i) enabling the student to familiarize with the technologies; (ii) setting-up a first use case for experimentations. Primarily, we will investigate whether LLMs can be used to interpret and understand EEG data. Once this is done, these data could be used to interpret human intentions and thoughts from EEG data. Subsequently, this important information can be leveraged to control a robot via a data-driven algorithm. We will set-up an experiment where a simple EEG headset is used to control a unicycle robot in unstructured environments. We will assess if the data from the EEG headset can be used to move the robot and if intelligent, data-driven, algorithms can be used to make the robot interact with the person.

O3 – Collaborative Tasks. With this objective, we will expand upon the architecture from O2 this time to include feedback from the robot to the person. We will target a collaborative task (using eventually a virtual reality set-up) that will require state-of-the-art EEG headsets. At this stage, we expect that the EEG might be complemented with other human modalities, i.e., movement or eye tracking to predict human intentions or action decisions for human-robot interaction. We envisage that this technology might be particularly relevant to people without language capabilities due to neural injury or disease.

O4 – architecture design. This objective is transversal to the previous ones. The nature of the project is such that, for each of the above objectives, we will need to set-up novel experiments, leveraging state of the art technologies. We will start with defining a basic architecture tackling O2 and subsequently expanding the architecture to enable the interaction tasks from 03.

Workplan. The plan will be developed in incremental tasks and periodic meetings will be scheduled with the supervisors. First, the student will start with becoming familiar with the existing literature in the areas related to the project. The output of this first step will be the definition of the technological and methodological constraints that the architecture will need to satisfy. Then, in the second phase, the student will develop the methodology and perform deployments/tests on small-scale problems (as described in the project objectives). The final part of the project will see the student finalizing the architecture deploying it on the selected application.

See the list of references for further details on the different aspects of the project.

Relevance to the MERC PhD Program (max 2000 characters)

The project is, by its very definition, **interdisciplinary and ambitious**. We aim at designing *a new* architecture towards the next generation of human-robot interactions. The link with complexity is apparent. We expect the architecture itself to be a complex entity, which will need to interact with other agents across a web of (potentially, time-varying) connections.

See the list of references for further details on the different aspects of the project.

Key references

Auletta, F., Kallen, R.W., di Bernardo, M. et al. Predicting and understanding human action decisions during skillful joint-action using supervised machine learning and explainable-AI. Sci Rep 13, 4992 (2023). https://doi.org/10.1038/s41598-023-31807-1

Yajing Si, Fali Li, Keyi Duan, Qin Tao, Cunbo Li, Zehong Cao, Yangsong Zhang, Bharat Biswal, Peiyang Li, Dezhong Yao, Peng Xu, Predicting individual decision-making responses based on single-trial EEG, NeuroImage, Volume 206, 2020, 116333, https://doi.org/10.1016/j.neuroimage.2019.116333

Jonathan W. Kim, Ahmed Alaa and Danilo Bernardo, EEG-GPT: Exploring Capabilities of Large Language Models for EEG Classification and Interpretation, preprint, 2024, https://arxiv.org/abs/2401.18006

Ziyang Huang, Mei Wang, A review of electroencephalogram signal processing methods for brain-controlled robots, Cognitive Robotics, Volume 1, 2021, Pages 111-124, https://doi.org/10.1016/j.cogr.2021.07.001

Yahui Wang, Suihuai Yu, Ning Ma, Jinlei Wang, Zhigang Hu, Zhuo Liu, Jibo He, Prediction of product design decision Making: An investigation of eye movements and EEG features, Advanced Engineering Informatics, Volume 45, 2020, 101095, https://doi.org/10.1016/j.aei.2020.101095

Joint supervision arrangements

Meetings will be scheduled on an as-needed basis, in order to ensure the effective development of the project. As a minimum, supervisor(s) will meet students at least once a week.

Location and length of the study period abroad (min 12 months)

The candidate will be able to spend a research period (or research periods) at Prof. Richardson's lab at the School of Psychological Sciences of Macquarie University (Sidney, Australia). Other research visits at leading labs working at the forefront of the topics covered by the project can be arranged.

Any other useful information

The project is best suited for students with a preference towards rigorously experimenting mathematical ideas (setting-up and designing experiments, processing of data etc…) and with a background in signal processing, dynamical systems and control. For further details on the background students can contact their potential supervisors.

Please return this form via email by no later than 9th February 2024 to merc@ssmeridionale.it