SSM
Scuola Superiore Meridionale

# MERC PhD Project Proposal 2023/2024

## Title of the research project:

Grounding Learning and Control onto the Free Energy Principle from Computational Neuroscience

## Keywords (up to five)

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

## Supervisors (at least two from two different areas):


***Supervisor 1***: *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*

***Supervisor 2***: *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*

***Supervisor 3***: *Mario di Bernardo (mario.dibernardo@unina.it)*
*Department of Electrical Engineering and Information Technology, University of Naples Federico II and Scuola Superiore Meridionale*
https://sites.google.com/site/dibernardogroup/home
*Complex Systems, Learning, Nonlinear Dynamics*


## Project description (max 5000 characters)

*The vision and the context of the project.* Over the past few years, a popular paradigm to design autonomous agents has become that of feeding them with data, using Reinforcement Learning to learn policies to perform a given task. Successes of this data- driven approach include the results in e.g., playing Go, outracing the world champions of Gran Turismo, learning policies to control plasmas from high-fidelity simulators and beating expert pilots in drones' competitions. In this context, with this project we tackle the following **key research questions/challenges**:

(i) how can we design complex networks of data-driven, learning, agents able to collaborate in a distributed way to fulfil a complex task?
(ii) Could the agents also safely collaborate with humans? And would the policies be *human-compatible*?
(iii) Can we make policy learning and computation reliable and suitable for real-time applications?

Gathering inspiration from neuroscience and cognitive sciences, to address the above questions, with this project we will design agents that ground their decision-making mechanisms into the free-energy principle (see below). Thus, we aim at designing agents that make decisions based exploiting

what psychologists, neuroscientists and cognitive scientists believe are the key mechanisms responsible for our ability to efficiently compute policies that are suitable for collaboration tasks. The **results** will be relevant to all applications, where agents need to be designed so that they can reliably fulfil *physical* tasks withing unknown, non-stationary and stochastic environments, such as the unknown exploration environment in the proposed application.
*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. The objectives are **inherently ambitious** and PhD students, also guided by their supervisors, will have the opportunity to selectively focus on a subset of specific objectives, in accordance with their individual interests, available time and opportunities.

O1 –*Free Energy Enabled Control***.** Informally, the Free Energy Hypothesis (FEH) postulates that our adaptive behaviors are aimed at minimizing our *surprise* in the data coming from the environment. If FEH is true, this implies that our actions are sampled from a randomized policy computed by our brain. With this objective we will design autonomous agents that, following the FEH, are able to compute optimal policies straight from the data seen from the environment. Moreover, we want the policies to be suitable for real-time applications. In general, it is known that minimizing the surprise in the data can easily become computationally intractable. Hence, we will explore the possibility of computing the agent policy by minimizing a cost functional that upper bounds the surprise. It is expected that this upper bound, following some prior work from one of the proponents, consists of an expected cost regularized with a Kullback-Leibler (KL) divergence term. We will then tackle the resulting optimal control problem by also unveiling and exploiting its rich connections with the exciting field of optimal transport and algorithms/ML methods inspired by statistical mechanics (Simulated Annealing, Restricted Boltzmann Machines). To early benchmark the results, we will use our policies on a simulated use-case that involves controlling a rover and a drone moving in an unstructured environment, while avoiding obstacles.

*O2 – Complex Networks of Free-Energy-Minimizing Agents to Control Networks.* With this objective, we will investigate what happens when the agents designed in O1 become nodes of a complex network and these nodes need to collaborate to perform a joint task. For concreteness, as a reference use case, building on the previous objective, we will consider a network of rovers, now equipped with robotic arms that need to carry certain (a-priori unknown objects). In general, effective collaborative strategies often arise when the agents know who they have in front of them, together with their intentions. With this objective, we will investigate how to mathematically formalize this observation, aiming at showing stability and optimality of the network behavior. We will thus consider the following three key aspects: (i) if the agents make decisions according to the policies from O1, what information do they need to exchange to make a collective decision? (ii) can the intentions/cost of an agent be inferred by the other agents in the network? Again, we will benchmark the results by considering the use case from O1, this time with a network of rovers and drones we will use our policies on a simulated use-case that involves controlling a rover and drones that need to safely navigate while carrying out a given object.

O3 – *From theory to the real world (and back)*. Our ambition is to ultimately deploy our results on agents with limited computational capabilities. While the application can be tailored towards the specific interests of the students, we plan to validate our results on the use-case from O1 and O2. Within the use case, the agents have no prior information about the environment where they are moving and they will need to *build* a simple structure. This scenario mirrors a challenging frontier application in the context of space exploration, where agents will need to build complex structures in environments that are, by their very own definition, largely unknown (i.e. the agents cannot be trained in the environment where the robot will operate), nonstationary, stochastic and nonlinear. Depending on time and opportunity, we will explore the possibility of deploying the algorithms onto a few real robotic agents.

*Workplan.* The project offers a mix of theory and experimentations and the workplan, tailored towards students' background, will be built to reflect this aspect giving to the project either a more methodological- or application-oriented angle: given the high ambition of each individual objective, students will have the opportunity to selectively focus on a subset of the objectives. 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 a preliminary methodology and a refined application. 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 methods and tools deploying them on the selected application. A simplified simulator of a rover moving in an unstructured environment is available to facilitate students towards familiarizing with the application/methodological concepts that will be developed. Within the project, students will also benefit from visiting labs at the forefront of the topics covered in the above objectives (see below for potential destinations/collaborations).

**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 *better* autonomous agents contaminating control, learning and optimization with ideas from other fields as outlined above. Depending on the students' interests, the main outcomes of the project will be a combination of novel methodological and application-oriented results. From the methodological viewpoint, the project will lead to a novel approach to synthesize policies from data. This will allow to tackle an application (as described above) that is difficult, if not impossible with classical control and learning methods. The link with complexity and risk is apparent. We expect the agent itself to be a complex entity, which will need to interact with other agents across a web of (potentially, time-varying) connections. As underpinned by the project objectives, jointly, the agent of this network will need to achieve safe collaboration, efficient use of data and handling the uncertainty from the environment.

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

## Key references

T. Appenzeller, "The AI revolution in Science," 2018. [Online]. Available: https://www.sciencemag.org/news/2017/07/ai-revolution-science

E. Horvitz, "AI, people and Society," Science, vol. 357, p. 7, 2017.

R. S. Sutton and A. G. Barto, Introduction to Reinforcement Learning (1st ed.), Cambridge, MA, USA: MIT Press, 1998.

P.R. Wurman, S. Barrett, K. Kawamoto, et al. "Outracing champion Gran Turismo drivers with deep reinforcement learning", Nature 602, 223–228, 2022

Degrave, J., Felici, F., Buchli, J. et al. "Magnetic control of tokamak plasmas through deep reinforcement learning", Nature 602, 414–419, 2022

E. Garrabé, G. Russo, Probabilistic design of optimal sequential decision-making algorithms in learning and control, Annual Reviews in Control, vol. 54, pp. 81-102, 2022

D. Gagliardi, G. Russo, "On a probabilistic approach to synthesize control policies from example datasets", Automatica, vol. 137, 110121, 2022

E. Garrabé, G. Russo, "On the design of autonomous agents from multiple data sources", IEEE Control Systems Letters, vol. 6, pp. 698 - 703, 2022

G. Russo, "On the crowdsourcing of behaviors for autonomous agents", IEEE Control Systems Letters, vol. 5, pp. 1321-1326, 2021

B. M. Lake et al., “Human-level concept learning through probabilistic program induction,” Science, vol. 350, no. 6266, pp. 1332–1338, 2015

E. Crisostomi et al., Analytics for the sharing economy: Mathematics, Engineering and Business perspectives. Springer, 2020

Stefano V. Albrecht and Peter Stone. Autonomous Agents Modelling Other Agents: A Comprehensive Survey and Open Problems. Artificial Intelligence. Volume 258. 2018.

E. Garrabé, G. Russo, "CRAWLING: a Crowdsourcing Algorithm on Wheels for Smart Parking", Scientific Reports, vol. 13, 16617, 2023 Collection in Advanced Transportation Systems

E. Garrabé, M. Lamberti, G. Russo, "Optimal Decision-Making for Autonomous Agents via Data Composition", IEEE Control Systems Letters, vol. 7, pp. 2557-2562, 2023

Lucian Busoniu, Robert Babuska, Bart De Schutter. A Comprehensive Survey of Multiagent Reinforcement Learning. In IEEE Transactions on Systems, Man and Cybernetics. Volume 38. Issue 2. March 2008.

Karl Tuyls and Gerhard Weiss. Multiagent Learning: Basics, Challenges and Prospects. AI Magazine. Volume 33. Issue 3. 2012.

Karl Tuylis and Peter Stone. Multiagent Learning Paradigms. In Francesco Belardinelli and Estefania Argente, editors, Multi-agent Systems and Agreement Technologies. Lecture Notes in Artificial Intelligence. Pages 3-21. Springer 2018.

R. M. Murray, K. J. Astrom, S. P. Boyd, R. W. Brockett, G. Stein. Future directions in control in an information-rich world. IEEE Control Systems Magazine. Volume 23. March 2003.

K. Friston, "The free-energy principle: a rough guide to the brain?," Cell, vol. 13, pp. 293-301, 2009

## 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. Specifically, within this project, students will also have the opportunity to spend time abroad at one or more of the following potential destinations:*

*(i) the group of Prof. Florian Dörfler at ETH Zurich;*
*(ii) the group of Professor Robert Shorten at the Dyson School of Design Engineering within Imperial College London.*

## Any other useful information

*While the project can be tailored towards students with more application-oriented interests, the project is best suited for students with a preference towards mathematical rigour (a proof-oriented mindset) and with a background in dynamical systems, control and learning. 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***