SSM
Scuola Superiore Meridionale

MERC PhD Project Proposal 2024/2025

Title of the research project:

Building Machines That Play Hard And Work Hard (as a team)

Keywords (up to five)

Autonomous agents, Multi-Agent Learning & Control, Complex Systems, Optimization, Decision-making

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, Nonlinear/Optimal Control, Complex Systems, Optimization

***Supervisor 2**: Mirco Musolesi (m.musolesi@ucl.ac.uk)*
Dept. of Computer Science at University College London, MERC PhD Board Member
www.mircomusolesi.org
Human-centred Artificial Intelligence, Reinforcement Learning, Modelling of social systems (human, artificial, hybrid)

Project description (max 5000 characters)

<u>*The vision and the context.*</u> The title of the project evokes its research vision: building machines – autonomous agents – that reliably perform complex tasks. The tasks require physical interactions with other machines, humans and challenging environments. Consequently, the **overarching goal** of the project is to design the methodological and computational tools to tackle some of the key bottlenecks towards reliable deployment of autonomous agents in real-world settings, including:

(i) Reliably making optimal decisions in non-stationary environments filled with other machines – and humans;
(ii) Make the policies interpretable, so that they can be understood by other agents – natural or artificial;
(iii) Predict the intentions of the other agents – and acting upon these predictions, which by their very definition, are wrong to some extent;
(iv) Achieve all the above while guaranteeing stability in nonlinear and stochastic environments.

This ambitious project is interdisciplinary – and it requires an interdisciplinary approach. Namely, holistically integrating concepts from complex systems, machine learning, dynamical systems, and control theory, our technical approach is threefold:

(i) Design agent-centric decision-making mechanisms for policy computation. These mechanisms, housed on an agent, implement *local* learning/control rules. By local, we mean that each agent only receives information from the nearby environment, from neighbouring agents and from humans – thus creating a complex cyber-physical network system. The design goal is than that of guaranteeing a desired global emerging behavior;

(ii) Ensuring, in the design process, that the policy is interpretable in the sense that each agent can understand – and hence predict – the intention of its neighbours. Non-cooperative/byzantine behaviors will be taken into account at this stage and the policies will be designed so that they are resilient to these misbehaving agents. Also, the machines we want to build will also be equipped with the ability of understanding the intentions of the humans that are within their environment;

(iii) leverage the brand new CoRE (Complexity and Risk Engineering) Lab at the Scuola Superiore Meridionale to bring our multi-agent systems to life on a timely – and extremely relevant – application that involves designing a distributed intelligence to control a complex network of drones and rovers competing against each other in the popular *capture the flag* game.

The results of this highly interdisciplinary project will be transversal to several communities, including systems and control engineering, and machine learning. The project also builds on a solid scientific foundation (see also list of references).

<u>*Project objectives and methodology.*</u> The detailed research program of this ambitious project will be shaped based on the **interests of students**. A preliminary list of concrete research objectives is given below.

O1 – *Decision-Making and Network Coordination***.** This objective is concerned with the design of distributed computational models that output adaptive control/learning policies for real-time applications. In essence the model implements a distributed and resilient mechanism for learning and control in settings where multiple autonomous agents must collaborate – despite malicious agents – to achieve shared/distributed goals. The computational model will be housed on each of the agents and the resulting policies guarantee: (i) the emergence of the desired closed-loop emerging behaviour in all operating condition; (ii) robustness against environmental variability induced by other agents and humans in the environment. Moreover, depending on time and opportunity, this objective also investigates what are the limits of this architecture, with emphasis put on important questions such as: *how many malicious agents can our model support? What is the best performance that the model can obtain when there are non-collaborative agents?* The tools to achieve this might draw inspiration from the rich literature on distributed control and optimization, building on some of the proponents' recent works. **In summary**, the methodological/computational tools from O1 yield: 1) an Edge Policy Engine (EPE) – a lightweight distributed computational model that, hosted on the agents, makes them part of a complex cyber-physical system and guarantees task fulfilment in a fully decentralized way; 2) its performance limits.

O2 – Embedding Humans-in-the-loop. Given the EPE from O1, in this objective human-agents interactions are explicitly introduced into our computational model. This allows to tackle applications where machines need to physically share space with humans in unstructured environments. Here, the focus shifts to designing systems capable of understanding, predicting, and reacting to human intent, while preserving safety and ensuring policy interpretability. This involves building models that can anticipate human actions and integrating them into the agent's control/learning loop. The proposers have experimented with this concept in the past and this experience will be leveraged. The methodological backbone may draw inspiration from game theory (e.g., leveraging the formalism from Shafiei et. al. from the references) and optimal transports (leveraging the tools from Centorrino et. al. from the references) – two exciting technical areas at the intersection of control and machine learning. A critical goal is to ensure the agent's behaviour remains transparent: to this aim, we ensure that the policy is interpretable and, building on e.g., maximum likelihood arguments, we obtain online optimization tools to infer the cost, and hence intentions, driving the actions of the agents. **In summary**, O2 outputs the computational and methodological tools to endow our agents with the ability of actively interacting with humans, predicting and acting upon their intentions.

O3 – Embodiment. With this objective, we aim at embodying the computational models from O1 and O2 into robotic agents. Equipped with our model, each agent is able to process heterogenous streams of data and reliably implement a distributed intelligence to fulfil a shared goal against competing entities/agents. This part of the work strategically leverages the robots available at the recently established CoRE Lab (crazyflies drones, turtlebots, duckiebots all sharing the same physical space in an *arena*). **In summary**, O3 brings the results from O1 and O2 to life using in-house equipment and thus providing hands-on experience, translating theory to real-world systems.

O4 – *Experiments*. Our computational model will be benchmarked, tested and validated on challenging and exciting applications. While the applications can be tailored towards the students' interests, also taking into account opportunity and time, the plan is to deploy our model on two teams of rovers and drones headed by a person (the captain). The two teams will compete against each other in the popular *capture the flag game*. In this game, a point is scored when one of the team members captures a flag (in this case, an electronic marker in the arena) from the other team and returns to its home base. The application has been preliminarily chosen not only because it is fun and pushes the limits of control and learning, but also because it is exemplary of all these applications across science and technology where complex organizations compete with each other and, to outperform the other, each of the member of the organization can take-on new roles, cover actions of team mates and – in genera – think strategically. This strategic thinking – common from simple bacteria to humans – is a salient feature that state-of-the-art autonomous agents simply lack! **In summary**, O4 brings the results of the project to an extremely timely – and relevant – application that not only pushes the limits of learning control, but it is also paradigmatic!

Workplan. The project 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 that has the potential to tackle the above objectives. Then, in the second phase, the student will develop the methodology and deploy/test the theory on small-scale problems. The final part of the project will see the student applying the methods and tools developed to a selection of real-world applications (see above) that are relevant to MERC.

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, highly interdisciplinary. The ambition of the project is to disrupt the way multi-agent systems are designed and analysed. Our aims are clearly related to modern intelligent complex systems. The project also has a link with risk, not only through the application, but also through the fact that the distributed intelligence we envision will need to safely operate in environments that are noisy.

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

Key references

Katherine M. Collins, Ilia Sucholutsky, Umang Bhatt, Kartik Chandra, Lionel Wong, Mina Lee, Cedegao E. Zhang, Tan Zhi-Xuan, Mark Ho, Vikash Mansinghka, Adrian Weller, Joshua B. Tenenbaum, and Thomas L. Griffiths, "Building machines that learn and think with people," Nature Human Behaviour, 8(10):1851–1863, 2024

B. Lake, T. Ullman, J. Tenenbaum and S. Gershman, "Building machines that learn and think like people," Behavioral and brain sciences, vol. E253, pp. 1-72, 2017

S. Xie, G. Russo, R. Middleton, "Scalability in nonlinear network systems affected by delays and disturbances", IEEE Transactions on Control of Network Systems, vol. 8, pp. 1128-1138, 2021

D. Fiore, G. Russo, "Resilient consensus for multi-agent systems subject to differential privacy requirements", Automatica, vol. 106, pp. 18-26, 2019

E. Garrabé, H. Jesawada, C. Del Vecchio, G. Russo, "On Convex Data-Driven Inverse Optimal Control for Nonlinear Stochastic Non-stationary Systems", Automatica, vol. 173, 112015, 2025

A. Shafiei, H. Jesawada, K. Friston, G. Russo, "Robust Decision-Making Via Free Energy Minimization", 2025, Submitted to Nature Communications.

V. Centorrino, F. Rossi, F. Bullo, G. Russo, "Proximal Gradient Dynamics and Feedback Control for Equality-Constrained Composite Optimization", 2025 (available on arXiv)

N. Anastassacos, S. Hailes, M. Musolesi. Partner Selection for the Emergence of Cooperation in Multi-agent Systems using Reinforcement Learning. Proceedings of AAAI'20. 2020.

N. Anastassacos, J. Garcia, S. Hailes, M. Musolesi. Cooperation and Reputation Dynamics with Reinforcement Learning. In Proceedings of AAMAS'21. 2021.

E. Tennant, S. Hailes, M. Musolesi. Modeling Moral Choices in Social Dilemmas with Multi-Agent Reinforcement Learning. In IJCAI'23. 2023.

G. Minelli, M. Musolesi. CoMIX: A Multi-agent Reinforcement Learning Training Architecture for Efficient Decentralized Coordination and Independent Decision Making. In Transactions on Machine Learning Research. 2024.

H. Long Fung. V.-A. Darvariu, S. Hailes and M. Musolesi. Trust-based Consensus in Multi-Agent Reinforcement Learning Systems. In Proceedings of RLC'24. 2024.

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

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.

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. Musolesi's lab at the University College London, UK. Moreover, while the expertise to carry-on the experiments is available at the CoRE Lab, if time and opportunity allows, the student will have the opportunity to spend research periods with a more application-oriented in other leading institutions, such as within Prof. Shorten's lab at Imperial College London (UK). Additionally, depending on the interests of the student, visits to other labs within the professional networks of the proposers can be accommodated.

Any other useful information

The project is best suited for ***ambitious students*** *with a preference towards students that enjoy translating formal deep mathematical thinking into real algorithms deployable on physical objects. For further details on the background students can contact their potential supervisors.*