

Thesis subject: Multi-objective and multi-agent reinforcement learning for the co-construction of ethical behaviors

Keywords

reinforcement learning ; multi-objective learning; multi-agent system; machine ethics ; explainable AI

Scientific context

With the recent, increasing progresses of Artificial Intelligence (AI), and seeing that “algorithms initially developed in the lab are increasingly being improved and deployed in society” [1], there is a crucial and pressing matter of ensuring that AI systems are aligned with (moral) values that are important to us (humans). These systems, by interacting with humans and, more generally, being immersed in our societies, have an impact on our lives. This urges AI researchers to develop more ethically-capable agents, shifting from ethics in design [2] to ethics by design with “explicit ethical agents” [15] able to behave ethically thanks to the integration of reasoning on and learning of ethics.

In this field of Machine Ethics [3], the Acceler-AI ANR project aims to develop an hybrid AI system able to learn behaviors aligned with moral values or ethical considerations, in co-construction with humans (regulators, users, stakeholders, etc.). The project adopts a human-centric approach, coupling lifelong-learning and non-invasive continuous human-AI system interaction to co-construct ethics, and integrating a normative regulation process that allows bounding the ethics learning process. The project is driven by a pluri-disciplinary consortium, with expertise in reinforcement learning, normative systems, human-computer interaction (HCI), and philosophy and ethics of science.

Two prototypical demonstrators will be developed targeting two different domains: energy based on a smart grid simulator¹ [4] and mobility relying on the “[Plateforme Territoire](#)”.

Thesis objectives

In this context, the objective of the thesis is the development of the system's lifelong learning of ethical behaviors, within controlled boundaries, integrated in a co-construction process with human feedback to guide the learning. The learning system will be composed of several interacting agents [5]: each agent is responsible for controlling/recommending actions based on moral preferences of the user with which it interacts ; and each agent also interacts with the other agents of the system (e.g. those interacting with other users of the transportation system). Thus **the thesis focuses on the combination of multi-objective and multi-agent reinforcement learning (MOMARL)**, to consider the decision of other learning agents and the multiple moral values (objectives) of the users. Because agents can learn undesirable behaviors, regulation mechanisms will be employed to ensure that agents operate within specified ethical boundaries while being sufficiently autonomous to learn ethics and adapt in response to evolution of the context and objectives. These mechanisms based on environment feedback (i.e. regimentation) or normative framework (i.e. enforcement) are used to constrain the RL of the ethical behaviors [27].

1 <https://github.com/ethicsai/ethical-smart-grid>

Very few works have tackled both MORL and MARL [6] and no work in the Machine Ethics field considers multi-objective, multi-agent [6-7], and human-centered approach [16]. The main contribution of the thesis would be to **propose a multi-objective multi-agent learning algorithm, able to identify sets of optimal policies, considering different trade-offs for the conflicting objectives and multiple agents, while operating within specified boundaries.** A proof-of-concept approach, developed in a previous project [4], is able to co-identify conflicting objectives and possible trade-offs between an artificial agent and a human user. A first approach in this thesis could be to extend this work to consider multiple agents in the same environment, and address trade-offs that could involve more than one agent, by proposing joint actions instead of unilateral actions. Most MARL approaches [8-9] require sharing information with other agents (Centralized Learning Decentralized Execution paradigm), which impairs privacy. In addition to the ‘by design’ ethical considerations, by improving the number and the quality of the found trade-offs with a joint policy, the contribution will have to abide by ‘in design’ ethical considerations. One such important aspect will be to preserve privacy during the data sharing among agents. For this part, intrinsically motivated social learning [13-14] will be considered.

Another aspect concerns the “dilemmas” situations, where several moral values are in conflict, and no single decision allows satisfying all of them at the same time: each choice will lead to regret. We argue that these situations cannot be “autonomously” settled by machines only, at least not how humans would like (expect) them to be settled. Thus another contribution of the thesis will be to propose an intelligible MOMARL approach taking into account several (more than three) objectives, and to be able to identify and settle situations of dilemmas, especially those requiring human intervention. First, as the number of “dilemmas” situations could be too high to be efficiently presented to end users, an exploration guided process based on intrinsically motivated reinforcement learning (e.g., curiosity models, learning progress,...) [10] will be investigated. Then, to allow the intelligible presentation of the alternatives to users, a refinement process will also be studied to classify dilemmas and involve the users through non-invasive HCI. We also propose to leverage human preferences to decide how to settle some dilemmas. To this end, the system must be able to use human feedback as a reward [11] or to learn models of users preferences, investigating approaches that learn preferences/profiles with few/no a priori data [12] and then adapt them through non-invasive HCI.

Work environment

The PhD student will join the [SyCoSMA](#) team of the [LIRIS](#) laboratory, located on the Doua of Lyon 1 University. There will be regular travel to [EMSE](#) (Saint-Etienne) and frequent interactions with participants of the project, especially those who are in charge of developing models and mechanisms for bounding the learning of ethics, and designing HCI interface.

The PhD student will be fully integrated to the Acceler-AI ANR project and will have to :

- participate to project meetings
- co-supervise interns in the project scope
- contribute to the writing of scientific articles
- participate to promoting and dissemination actions of achieved results (seminars, conferences, ...)
- possibility to realize casual teaching

Supervisors:

- [Laëtitia Matignon](#) (LIRIS - UCBL1)
- [Luis Gustavo Nardin](#) (LIMOS - EMSE)
- [Mathieu Guillermin](#) (UCLy)

Required profile:

- Computer science Master
- Skills in artificial intelligence
- Interests for philosophy of science and AI ethics
- Good level in english

Application process: Send CV, letter of motivation, academic transcript of the last 3 years, references or letters of recommendation ; by mail to laetitia.matignon@univ-lyon1.fr

Bibliography

- [1] Alexandra Luccioni and Yoshua Bengio, ‘On the morality of artificial intelligence’, arXiv preprint arXiv:1912.11945, (2019)
- [2] Dignum, V. Responsible artificial intelligence: how to develop and use AI in a responsible way. Springer Nature. (2019).
- [3] Michael Anderson and Susan Leigh Anderson, Machine ethics, Cambridge University Press, 2011.
- [4] Chaput Rémy, Learning behaviours aligned with moral values in a multi-agent system: guiding reinforcement learning with symbolic judgments, PhD Dissertation, 2022.
- [5] Chaput, R., Duval, J., Boissier, O., Guillermin, M., & Hassas, S. (2021). A Multi-Agent Approach to Combine Reasoning and Learning for an Ethical Behavior. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (pp. 13-23).
- [6] Rădulescu, Roxana, Patrick Mannion, Diederik M. Roijers, and Ann Nowé. “Multi-Objective Multi-Agent Decision Making: A Utility-Based Analysis and Survey.” *Autonomous Agents and Multi-Agent Systems* 34, no. 1 (December 9, 2019): 10.
- [7] Hayes, Conor F., Roxana Rădulescu, Eugenio Bargiacchi, Johan Källström, Matthew Macfarlane, Mathieu Reymond, Timothy Verstraeten, et al. “A Practical Guide to Multi-Objective Reinforcement Learning and Planning.” ArXiv:2103.09568 [Cs], March 17, 2021.
- [8] Panait, L., & Luke, S. (2005). Cooperative Multi-Agent Learning: The State of the Art. *Autonomous Agents and Multi-Agent Systems*, 11(3), 387–434.
- [9] Gronauer, Sven and Diepold, Klaus ; Multi-agent deep reinforcement learning: a survey, in *Artificial Intelligence Review*, 2022
- [10] Arthur Aubret, Laëtitia Matignon, Salima Hassas: An Information-Theoretic Perspective on Intrinsic Motivation in Reinforcement Learning: A Survey. *Entropy* 25(2): 327 (2023)
- [11] Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., and Amodei, D. (2017). Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems*, pages 4299–4307.
- [12] Strouse, D. J., McKee, K., Botvinick, M., Hughes, E., & Everett, R. (2021). Collaborating with humans without human data. *Advances in Neural Information Processing Systems*, 34.
- [13] Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P., Strouse, D. J., ... & De Freitas, N. (2019, May). Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In *International Conference on Machine Learning* (pp. 3040-3049). PMLR.

- [14] T. van der Heiden, C. Salge, E. Gavves, H. van Hoof , Robust Multi-Agent Reinforcement Learning with Social Empowerment for Coordination and Communication, 2020, arXiv:2012.08255 [cs].
- [15] Moor, J. H. “The Nature, Importance, and Difficulty of Machine Ethics.” IEEE Intelligent Systems 21, no. 4 (July 2006): 18–21. <https://doi.org/10.1109/MIS.2006.80>.
- [16] Nallur, Vivek. “Landscape of Machine Implemented Ethics.” Science and Engineering Ethics 26, no. 5 (October 1, 2020): 2381–99. <https://doi.org/10.1007/s11948-020-00236-y>.
- [27] Boella, G., Van Der Torre, L., & Verhagen, H. (2006). Introduction to normative multiagent systems. Computational & Mathematical Organization Theory, 12(2), 71-79