Universidade de Pernambuco
Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Dissertação de Mestrado

Área: Inteligência Computacional

Título: Dynamic Authority Allocation in Reinforcement Learning

Orientador – Bruno José Torres Fernandes (bjtf@ecomp.poli.br)
Coorientador – Davib Barrientos (djbr@ecomp.poli.br)

Descrição – Recent advances in reinforcement learning (RL) have enabled the control of increasingly complex articulated systems, achieving strong performance in manipulation, locomotion, and sequential decision-making tasks [1]. These advances have largely been driven by centralized learning paradigms, in which a single global policy observes the full system state and directly controls all degrees of freedom. While such approaches have proven effective in controlled and simulated environments, fully centralized RL architectures often face challenges related to scalability, robustness, and generalization as system dimensionality and task complexity increase [2].

In many practical control problems, articulated systems are composed of multiple joints or functional units that operate under different local constraints and at different temporal scales. Treating all joints as equally dependent on a global controller can lead to inefficient learning, increased coordination overhead, and brittle behaviors when the system is exposed to disturbances or partial failures [3]. Furthermore, centralized policies frequently suffer from credit assignment difficulties and reduced sample efficiency when required to coordinate fine-grained joint-level actions over long time horizons [4].

To mitigate these limitations, a variety of hierarchical and decentralized control approaches have been proposed, including hierarchical reinforcement learning, options-based methods, feudal reinforcement learning, and multi-agent reinforcement learning frameworks [4][5]. These methods decompose control either across multiple levels of abstraction or across interacting decision-making units. However, in most existing approaches, the decomposition of control authority is predefined and remains fixed throughout learning and execution, with the assignment of autonomy to specific joints or sub-tasks determined a priori [6].

In real-world scenarios, however, the need for global coordination is often highly context-dependent. Certain joints or sub-tasks may be effectively handled by local controllers under nominal operating conditions, while requiring global intervention only in situations involving strong inter-joint coupling, increased uncertainty, or task transitions. Fixed autonomy assumptions are unable to capture this variability, potentially resulting in unnecessary global control in simple situations or insufficient coordination when complex interactions arise [7]. This observation raises a fundamental research question: under what conditions should a sub-task operate autonomously, and when should it defer control to a global policy?

This project proposes to investigate dynamic authority allocation in reinforcement learning, in which control responsibility is not statically assigned but instead emerges through learning and interaction with the environment. In the proposed paradigm, local level controllers are capable of operating independently based on local state information, while retaining the ability to consult a global RL controller when specific conditions are met. Importantly, consultation is treated as a deliberate and explicit decision, rather than an exceptional or failure-driven event, allowing for selective, sparse, and context-aware global intervention.

From a theoretical standpoint, this approach aligns with recent perspectives that view reinforcement learning not only as a mechanism for action selection, but also as a framework for learning decision hierarchies and information flow within complex systems [8]. By explicitly modeling consultation as part of the control process, potentially with an associated cost, the proposed framework relates to ideas from hierarchical reinforcement learning, resource-rational decision-making, and adaptive control [9]. Unlike traditional hierarchical methods, this work does not assume that an optimal decomposition of control authority is known in advance, but instead seeks to learn when coordination is necessary and when local autonomy is sufficient.

The research will be conducted through a comparative study of three learning architectures: a fully centralized RL baseline, in which a global policy controls all joints at all times; a fixed autonomy baseline, in which a predefined subset of joints operates exclusively under local control; and a dynamic authority allocation framework, in which joints learn when to act locally and when to request guidance from the global controller. This comparative structure enables a principled analysis of the trade-offs between centralized control, static decomposition, and dynamically learned coordination strategies.

The central hypothesis of this work is that enabling joints to selectively consult a global controller can improve learning efficiency, robustness to perturbations, and generalization across tasks, while reducing unnecessary coordination and control complexity. By analyzing consultation patterns, joint-level autonomy, and task performance under varying conditions, the project aims to identify structural principles governing effective control decomposition in reinforcement learning systems.

In summary, this project seeks to develop, implement, and evaluate a reinforcement learning framework that explicitly studies how control authority should be distributed and dynamically adjusted across joints and tasks. Rather than proposing a single optimal controller, the research aims to contribute a conceptual and experimental foundation for understanding when autonomy is sufficient and when coordination is essential. The expected contributions include insights into scalable reinforcement learning architectures, adaptive hierarchy formation, and principled criteria for global intervention in complex articulated control systems.

Referências Bibliográficas

1. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *nature*, *518*(7540), 529-533.
2. Li, Y. (2017). Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*.
3. Pateria, S., Subagdja, B., Tan, A. H., & Quek, C. (2021). Hierarchical reinforcement learning: A comprehensive survey. *ACM Computing Surveys (CSUR)*, *54*(5), 1-35.
4. Sutton, R. S., & Barto, A. G. (1999). Reinforcement learning. *Journal of Cognitive Neuroscience*, *11*(1), 126-134.
5. Pateria, S., Subagdja, B., Tan, A. H., & Quek, C. (2021). Hierarchical reinforcement learning: A comprehensive survey. *ACM Computing Surveys (CSUR)*, *54*(5), 1-35.
6. Dayan, P. (1994). Feudal Q-learning. *1995*.
7. Vezhnevets, A. S., Osindero, S., Schaul, T., Heess, N., Jaderberg, M., Silver, D., & Kavukcuoglu, K. (2017, July). Feudal networks for hierarchical reinforcement learning. In *International conference on machine learning* (pp. 3540-3549). PMLR.
8. Hutsebaut-Buysse, M., Mets, K., & Latré, S. (2022). Hierarchical reinforcement learning: A survey and open research challenges. *Machine Learning and Knowledge Extraction*, *4*(1), 172-221.
9. Bacon, P. L., Harb, J., & Precup, D. (2017, February). The option-critic architecture. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 31, No. 1).