AdLeap-MAS: An Open-source Multi-Agent Simulator for Ad-hoc Reasoning

Demonstration Track

Matheus Aparecido do Carmo Alves
Lancaster University
United Kingdom
m.a.docarmoalves@lancaster.ac.uk

Amokh Varma
Indian Institute of Technology, Delhi
India
Amokh.Varma.mt618@maths.iitd.ac.in

Yehia Elkhatib
University of Glasgow
United Kingdom
yehia.elkhatib@glasgow.ac.uk

Leandro Soriano Marcolino
Lancaster University
United Kingdom
l.marcolino@lancaster.ac.uk

ABSTRACT

Ad-hoc reasoning models are recurrently used to solve some of our daily tasks. Intending to avoid worthless investments or spend valuable resources, these smart systems requires a proper evaluation before acting in the real-world. In this paper, we demonstrate *AdLeap-MAS*, a novel framework focused on enabling quick and easy testing of smart algorithms in ad-hoc reasoning domains.

KEYWORDS

Simulation Framework; Open-source; Ad-hoc Reasoning; Online Planning; Autonomous Systems.

ACM Reference Format:
Matheus Aparecido do Carmo Alves, Amokh Varma, Yehia Elkhatib, and Leandro Soriano Marcolino. 2022. AdLeap-MAS: An Open-source Multi-Agent Simulator for Ad-hoc Reasoning: Demonstration Track. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 3 pages.

1 INTRODUCTION

Autonomous systems play notable roles in contemporary society. Facing the increasing number of devices (agents) in the environment and the rising of problems with higher complexity, there is a need for new intelligence methods capable of solving tasks, learning about the context and handling uncertainties in an online-manner. A typical approach presented by the state-of-art is to assign multiple intelligent systems (agents) to solve a common objective and use an ad-hoc teamwork model to coordinate this Multi-Agent system [1, 10, 16, 18, 20, 23, 26]. However, they miss the opportunity to generalise their model by considering different roles in the environment (e.g., where agents can be potential teammates or opponents) – defining what we denominate as an *ad-hoc reasoning domain.*

On the other hand, building an intelligent system implies also testing and evaluating it. Simulators are important tools for the advancement of ad-hoc reasoning research [12]. However, current literature suggests that each researcher is implementing their own scenarios and using different custom-built simulators for similar purposes [2, 5, 7, 14, 15, 17, 18, 21, 25]. Therefore, a common platform that is capable of assembling different problems and algorithm implementations is still missing.

We propose the open-source *Adaptative Learning and Planning Multi-agent Simulator* (*AdLeap-MAS*), a novel framework focused on simulating ad-hoc reasoning problems, where potential types/policies for other agents are estimated, and sampled during an on-line decision-making process. We offer base classes for implementing new problems and algorithms, besides ready-to-use common benchmarks found in the literature. This proposal supports the execution of reactive algorithms, neural networks, estimation methods, reinforcement learning and on-line planning application over full and partial observability, only requiring the connection of algorithms to the ad-hoc reasoning model. In this way, our contributions can be summarised as: **(i)** first simulator that allows a quick switch of learning and planning algorithms across different ad-hoc reasoning scenarios; **(ii)** *AdLeap-MAS* enables the execution of multiple reasoning agents that run independently; **(iii)** our architecture guarantees information security while running scenarios under partial observability, *i.e.,* agents do not have access to any forbidden information, and; **(iv)** a standard set of benchmark algorithms and problems to allow fair and quick experiments.

2 DESIGN FEATURES

The *AdLeap-MAS*'s architecture is based on unilateral and cyclical module communication, where the information within the framework must be delivered or received directly and exclusively by one module from another in the architecture. Such design enables the problem simulation as a step-by-step process, processing each fragment of the simulation independently. The 3 main modules are – Environment, Decision-making, and Components modules. The Environment module is responsible for all the simulations and also makes sure that certain information is hidden from the Decision-making module, to ensure that partial observability is not violated. The Components module controls the different dynamic parts of the environment – such as agents, tasks to be completed etc. This allows the user to use the Environment module as a black-box, which takes the action as input and returns the observation, without revealing any extra information. The high-level overview of *AdLeap-MAS* can be seen in Figure 1.

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

3 PROBLEM SCENARIOS

Currently, *AdLeap-MAS* offers 6 different environments: three of which are ad-hoc reasoning domains (1.abc) while the other three are traditional partial observability toy-problems (2.abc).

(1.a) The Level-based Foraging (LBF) represents an ad-hoc teamwork domain where the ad-hoc agent tries to maximise the number of boxes collected by its team while learning the environment, teammate features and deciding its own actions [1].

(1.b) Capture the Prey (CTP) is a domain derived from the traditional "pursuit game" [3, 4], where a team of hunters tries to catch all the preys in the environment. They need to surround all the preys in order to capture them before the timer expires.

(1.c) The "Truco" Card Game (TCG) is based on the popular Brazilian card game, where the agents have a small window of observation to make decisions and maximise their chance of winning.

(2.a) The Maze (MZ) is an active localisation problem where the ad-hoc agent navigates a toroidal grid-world to gather the available observation and figure out its actual position [22].

(2.b) The Rock Sampling (RS) domain, in which a robot moves around a grid-world and tries to maximise the number of "good" rocks collected while exploring the unknown map [19].

(2.c) The Tiger Domain (TG) is a classic POMDP domain where the ad-hoc agent must open one of two doors: one has a tiger and the other a treasure. The agent can listen to the tiger or make a decision of which door to open without this observation [11].

Figure 2 illustrates the above domains. We provide an easy-to-use template for the implementation of a new environment, which is versatile as the figure shows. The template is available at the framework GitHub's page with more information about the environments and simulator usage. Additionally, *AdLeap-MAS* offers some ready-to-use baselines for planning experiments, besides state-of-art estimation methods and several other reactive methods.

4 RELATED WORK

Intending to point out the major differences between the *AdLeap-MAS* and the frameworks which are currently available for similar purposes, we discuss some state-of-art proposals in detail.

OpenSpiel is a reinforcement learning framework widely used for the evaluation of planning algorithms [13]. However, even though it presents a collection of environments and algorithms, OpenSpiel is not focused on the simulation of ad-hoc reasoning domains. Furthermore, *AdLeap-MAS* enables an easier swap of algorithms between environments and agents using its component-based architecture.

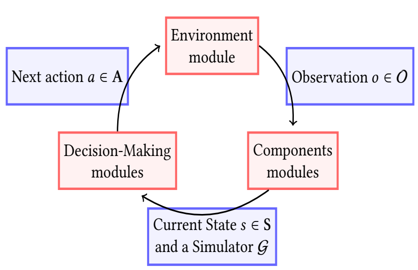

Figure 1: The *AdLeap-MAS* high-level workflow, indicating the information delivered at each step of the simulation.

Similar limitations emerge in the GAMA project [8], which focuses on the Multi-Agents context, but does not address ad-hoc reasoning domains. Despite the issue of acquiring a world model, the framework also requires an understanding of its dedicated programming language before utilisation.

In the literature, we also find simulators that are focused on tackling the representation and simulation of real robotics, such as Gazebo [12] and Stage [24]. Even having the ability to present high-fidelity simulations of multi-robot problems, these frameworks do not support learning/planning algorithms.

Finally, Open-AI Gym [6] is a Python package that provides a collection of benchmarks to run reinforcement learning tests by abstracting the environment. We extend the benefits of the Open-AI Gym platform and also improve its range of applications. By directly modelling and offering support of ad-hoc reasoning applications, we specialised the package for evaluation and simulation into the context. Our framework also handles modifying visibility restrictions without requiring further implementation.

5 ADLEAP-MAS READY-TO-GO

Our GitHub's page[1] furnishes the users with extensive documentation on how to use the framework. Moreover, we also release an introductory video[2] to demonstrate the functionalities of our simulator and facilitate understanding of its operation.

This paper represents our willingness to spread the current *AdLeap-MAS*'s results. We want to allow the development of research in a collaborative manner, capable of improving the overall results found by the community in the short and long term. As mentioned, our purpose is not to surpass the capabilities and functionalities of other simulators. Instead, our aim is to build a reliable solution that alleviates the difficulty of running experiments, and the complexity of fairly comparing algorithms for different problems without losing trust in the collected results.

Finally, we are continuing to work on *AdLeap-MAS* and improve its environments. Our current project is focused on developing a problem within the *continuous action and state spaces* – denominated Smart Fire Brigade Environment (based on [9]) – where we want to go towards more realistic scenarios, evaluating real-systems constraints, costs and application.

[1] ***AdLeap-MAS*'s GitHub:** https://github.com/lsmcolab/adleap-mas/
[2] **Introductory Video:** https://youtu.be/xCXFAyvofHo

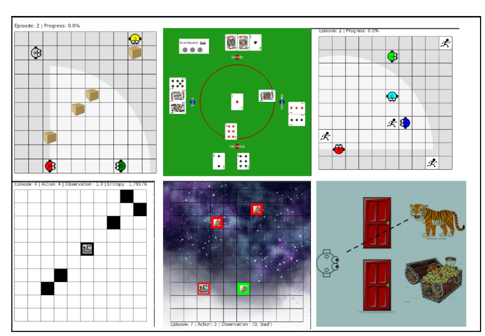

Figure 2: Different environments running in *AdLeap-MAS*: LBF, TCG, CTP, MZ, RS, and TG domain, respectively.

REFERENCES

[1] S. V. Albrecht and P. Stone. 2018. Autonomous Agents Modelling Other Agents: A Comprehensive Survey and Open Problems. *Artificial Intelligence (AIJ)* 258 (2018), 66–95.

[2] Mona Alshehri, Napoleon N Reyes, and Andre LC Barczak. 2020. Evolving Meta-Level Reasoning with Reinforcement Learning and A* for Coordinated Multi-Agent Path-planning.. In *AAMAS*. 1744–1746.

[3] Samuel Barrett and Peter Stone. 2012. An Analysis Framework for Ad Hoc Teamwork Tasks. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 1* (Valencia, Spain) *(AAMAS '12)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 357–364.

[4] Samuel Barrett, Peter Stone, and Sarit Kraus. 2011. Empirical Evaluation of Ad Hoc Teamwork in the Pursuit Domain. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*.

[5] Felix Brandt and Martin Bullinger. 2020. Finding and Recognizing Popular Coalition Structures.. In *AAMAS*. 195–203.

[6] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. 2016. Openai gym. *arXiv preprint arXiv:1606.01540* (2016).

[7] Giuseppe De Giacomo and Yves Lespérance. 2020. Goal Formation through Interaction in the Situation Calculus: A Formal Account Grounded in Behavioral Science.. In *AAMAS*. 294–302.

[8] Alexis Drogoul, Edouard Amouroux, Philippe Caillou, Benoit Gaudou, Arnaud Grignard, Nicolas Marilleau, Patrick Taillandier, Maroussia Vavasseur, Duc-An Vo, and Jean-Daniel Zucker. 2013. Gama: multi-level and complex environment for agent-based models and simulations. In *12th International Conference on Autonomous agents and multi-agent systems*. Ifaamas, 2–p.

[9] Maria Gini. 2017. Multi-robot allocation of tasks with temporal and ordering constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31.

[10] A. Guez, D. Silver, and P. Dayan. 2013. Scalable and Efficient Bayes-Adaptive Reinforcement Learning Based on Monte-Carlo Tree Search. *Journal of Artificial Intelligence Research (JAIR)* 48 (2013).

[11] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. 1998. Planning and Acting in Partially Observable Stochastic Domains. *Artif. Intell.* 101, 1–2 (may 1998), 99–134.

[12] Nathan Koenig and Andrew Howard. 2004. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, Vol. 3. IEEE, 2149–2154.

[13] Marc Lanctot, Edward Lockhart, Jean-Baptiste Lespiau, Vinicius Zambaldi, Satyaki Upadhyay, Julien Pérolat, Sriram Srinivasan, Finbarr Timbers, Karl Tuyls, Shayegan Omidshafiei, et al. 2019. OpenSpiel: A framework for reinforcement learning in games. *arXiv preprint arXiv:1908.09453* (2019).

[14] Błażej Osiński, Adam Jakubowski, Paweł Zięcina, Piotr Miłoś, Christopher Galias, Silviu Homoceanu, and Henryk Michalewski. 2020. Simulation-based reinforcement learning for real-world autonomous driving. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 6411–6418.

[15] Praveen Palanisamy. 2020. Multi-agent connected autonomous driving using deep reinforcement learning. In *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–7.

[16] Lukasz Pelcner, Shaling Li, Matheus Do Carmo Alves, Leandro Soriano Marcolino, and Alex Collins. 2020. Real-time learning and planning in environments with swarms: a hierarchical and a parameter-based simulation approach. (2020).

[17] Kambiz Rasoulkhani, Ali Mostafavi, Maria Presa Reyes, and Mostafa Batouli. 2020. Resilience planning in hazards-humans-infrastructure nexus: A multi-agent simulation for exploratory assessment of coastal water supply infrastructure adaptation to sea-level rise. *Environmental Modelling & Software* 125 (2020), 104636.

[18] Elnaz Shafipour Yourdshahi, Matheus Do Carmo Alves, Leandro Soriano Marcolino, and Plamen Angelov. 2020. Decentralised Task Allocation in the Fog: Estimators for Effective Ad-hoc Teamwork. In *11th International Workshop on Optimization and Learning in Multiagent Systems*.

[19] Trey Smith and Reid Simmons. 2004. Heuristic Search Value Iteration for POMDPs. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence* (Banff, Canada) *(UAI '04)*. AUAI Press, Arlington, Virginia, USA, 520–527.

[20] Peter Stone, Gal A. Kaminka, Sarit Kraus, and Jeffrey S. Rosenschein. 2010. Ad Hoc Autonomous Agent Teams: Collaboration without Pre-Coordination. In *Proceedings of the Twenty-Fourth Conference on Artificial Intelligence (AAAI)*.

[21] Khadija Tazi, Fouad Mohamed Abbou, and Farid Abdi. 2020. Multi-agent system for microgrids: design, optimization and performance. *Artificial Intelligence Review* 53, 2 (2020), 1233–1292.

[22] Vincent Thomas, Gérémy Hutin, and Olivier Buffet. 2021. Monte Carlo Information-Oriented Planning. arXiv:2103.11345 [cs.AI]

[23] Maulesh Trivedi and Prashant Doshi. 2018. Inverse Learning of Robot Behavior for Collaborative Planning. In *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

[24] Richard Vaughan. 2008. Massively multi-robot simulation in stage. *Swarm intelligence* 2, 2 (2008), 189–208.

[25] Kai Wang, Andrew Perrault, Aditya Mate, and Milind Tambe. 2020. Scalable Game-Focused Learning of Adversary Models: Data-to-Decisions in Network Security Games. In *AAMAS*. 1449–1457.

[26] Elnaz Shafipour Yourdshahi, Thomas Pinder, Gauri Dhawan, Leandro Soriano Marcolino, and Plamen Angelov. 2018. Towards Large Scale Ad-hoc Teamwork. In *2018 IEEE International Conference on Agents (ICA)*. IEEE, 44–49.