engineering proceedings

MDPI

Proceeding Paper

Enhancing Heterogeneous Multi-Robot Teaming for Planetary Exploration †

Amrita Suresh [1,*], Melvin Laux [1], Wiebke Brinkmann [2], Leon C. Danter [2] and Frank Kirchner [1,2]

[1] Faculty 3—Mathematics and Computer Science, University of Bremen, Robert-Hooke-Str. 1, 28359 Bremen, Germany; melvin.laux@uni-bremen.de (M.L.); frank.kirchner@informatik.uni-bremen.de (F.K.)
[2] German Research Center for Artificial Intelligence GmbH (DFKI), 67663 Kaiserslautern, Germany; wiebke.brinkmann@dfki.de (W.B.); leon.danter@dfki.de (L.C.D.)
* Correspondence: amrita@uni-bremen.de
† Presented at the 14th EASN International Conference on "Innovation in Aviation & Space towards sustainability today & tomorrow", Thessaloniki, Greece, 8–11 October 2024.

Abstract: Future space missions will include multi-robot systems, with greater autonomy and a large degree of heterogeneity for a wider range of task capabilities and redundancy. It is imperative that both software (learning models, parallelizing capabilities, resource distribution, etc.) and hardware factors must be considered during decentralized task negotiation to lead to better performance of the team. By utilizing the formalism of contextual Markov decision processes, team composition can be incorporated into the learning process and used for more meaningful and reliable evaluation using measures such as total time, overall consumed energy, performance feedback from tasks, or damage incurred. Improved team performance will in turn enhance the overall results of the mission. Planetary exploration tasks often involve time, communication and energy constraints. Such missions are also prone to noisy sensor data (e.g., camera images distorted by dust), as well as wear and tear on hardware (e.g., wheels, manipulators). To ensure that such factors do not jeopardize the mission, they must be taken into account. Therefore, this paper describes a software framework for the reliable execution of tasks in constrained and dynamic environments. Our work leverages the advantages of heterogeneity for more resilient planetary missions by addressing two aspects—first, the integration of hardware parameters into the negotiation process, and second the analysis of how the integration of team performance metrics, particularly adaptability and mutual support, in task negotiation plays a role in the overall mission success.

Keywords: heterogeneity; exploration; contextual RL; task negotiation; robotic teaming

check for updates

Academic Editors: Spiros Pantelakis, Andreas Strohmayer and Nikolaos Michailidis

Published: 8 May 2025

Citation: Suresh, A.; Laux, M.; Brinkmann, W.; Danter, L.C.; Kirchner, F. Enhancing Heterogeneous Multi-Robot Teaming for Planetary Exploration. *Eng. Proc.* **2025**, *90*, 112. https://doi.org/10.3390/engproc2025090112

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Planetary exploration is pivotal to expanding our understanding of the universe. In recent decades, there has been a surge in launch tests and unmanned missions to our celestial neighbours [1]. Despite the fact that these missions have contributed to significant scientific discovery, most have merely involved single-robot systems, except for NASA's dual deployment of Perseverance and Ingenuity [2]. These missions often suffer from confined task capability, limited redundancy, and reliance on partial tele-operation from Earth. This urges future missions to increasingly rely on multi-robot systems (MRSs) with heterogeneous members to handle diverse tasks and situations. The efficiency of a heterogeneous MRSs (HMRSs) greatly depends on effective task negotiation [3]. In such a decentralized approach, each robot autonomously acquires and executes tasks while adapting to real-time

conditions. Several algorithms exist for multi-agent task negotiation, coordination, and exploration [4]. However, most existing methods fall short when deployed in real-world applications due to their failure to account for critical hardware-related parameters, which can take a significant toll on mission performance.

Exploring planetary environments poses a multitude of challenges, such as sensor noise, wear and tear on mechanical components, environmental dynamics, communication limitations, and the finite lifespan of hardware. In practical scenarios, these hardware and environmental variables are often hard-coded into high-level controllers, limiting the flexibility and adaptability of robots in response to real-time issues. As a result, robotic systems are less autonomous and rely on partial tele-operation from Earth.

This paper proposes a novel solution approach to this problem by introducing contextual Markov Decision Processes (CMDPs) into the task negotiation process in a multi-robot system. Several hardware and environmental states can be encoded as contexts integrated into the robot's decision-making process. This allows for dynamic adaptation to configurational changes and environmental uncertainty, enabling more resilient and flexible task re-allocation. We have chosen an auction-based task negotiation algorithm. The robotic team is evaluated using two key terms that are inspired by human teaming concepts, namely, adaptability and mutual support. We present a novel approach to multi-robot task negotiation by:

- Incorporating hardware and environmental state as context into the learning process to boost mission resilience
- Employing a structured evaluation procedure fit to meaningfully test the proposed methodology and verify the statistical significance of experimental results

The remainder of the paper is structured as follows. Section 2 introduces the foundational concepts of contextual reinforcement learning (cRL) and CMDPs and their relevance to multi-agent systems. Next, Section 3 briefly reviews existing space robotic systems, heterogeneity, and context-based learning and highlights the limitations of previously deployed multi-robot systems for exploration in unpredictable environments. Section 4 presents the proposed method. Subsequently, Section 5 describes the metrics used to assess the proposed CMDP-based task negotiation framework, focusing on adaptability, mutual support within the robotic team, and overall team performance in dynamic environments. Finally, Section 6 presents potential future works and summarizes the potential benefits of the proposed method.

2. Background

The reinforcement learning (RL) problem is most commonly formalized using Markov Decision Processes (MDPs) [5]. An MDP describes a single RL task in the form of a tuple $\mathcal{T} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \mu_0)$, where $\mathcal{S}$ describes the state space, $\mathcal{A}$ the action space, $\mathcal{P}$ the transition dynamics, $\mathcal{R}$ the reward function, and μ_0 the initial state distribution [6]. However, while this formulation is sufficient for many academic or virtual applications, e.g., playing video games [7,8], most real-world tasks require the learning agent to behave appropriately in situations that differ from the one they were trained in. For example, a planetary exploration robot may encounter situations during deployment that could not be foreseen at the time of training, long before its arrival at the planet it is meant to explore. To precisely formalize the concept of generalization over tasks, the MDP framework can be extended by including the notion of context in contextual Markov decision processes (CMDPs) [9,10]. A CMDP allows us to accurately describe groups of different but related MDPs by defining a context space $\mathcal{C}$ and a mapping $\mathcal{M}$ from any given context $c \in \mathcal{C}$ to a context-dependent MDP $\mathcal{T}^c = (\mathcal{S}, \mathcal{A}, \mathcal{P}^c, \mathcal{R}^c, \mu_0^c)$. Note that state and action spaces remain constant across all tasks defined by the CMDP. A context space $\mathcal{C}$ can either be discrete set or a distribution from

which contexts can be sampled. This formulation context also allows for the concrete definition of test and training context sets [11]. The conceptual idea behind CMDPs is illustrated in Figure 1. By defining a set of training contexts $\mathcal{C}_{train} \subseteq \mathcal{C}$ and $\mathcal{C}_{test} \subseteq \mathcal{C}$, we can evaluate a policy's ability to generalize to unseen situations effectively. This formulation of generalization also more closely matches the training and evaluation protocols of traditional supervised learning, e.g., classification tasks. Moreover, such an evaluation procedure allows for a much more rigorous and systematic testing and verification of RL-based systems—a critical component of real-world deployment of autonomous systems [12].

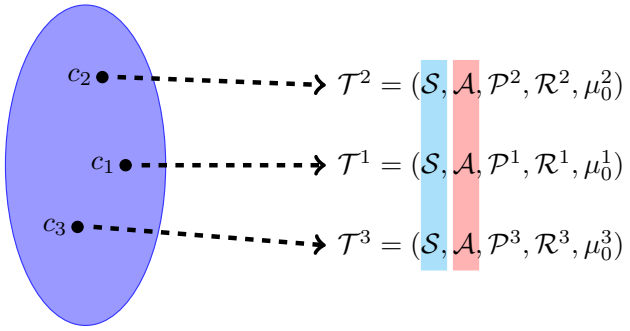


Figure 1. Illustration of CMDPs. Each context in the context space maps to an individual MDP. The state and action spaces remain constant across all MDPs, while transition probabilities, reward functions, and initial state distributions may vary.

3. State of the Art

3.1. Heterogeneous Multi-Robot Systems

Heterogeneous multi-robot systems generally consist of at least two distinct individual systems. At least one system is an active robotic system with wheels or legs or that is airborne. Tasks covered by such cooperation are, e.g., cooperative mapping and navigation, manipulation and transport, and modularity and reconfiguration for enhanced versatility, as well as autonomous repair, maintenance, and rescue. Task negotiation among robots is vital to a decentralized approach, wherein the robots evaluate their own capabilities, resources, and current state (e.g., battery level, proximity, or payload capacity) in relation to the team's goals and the task requirements and negotiate for suitable tasks. There are numerous algorithms to achieve this such as those discussed in [13–15].

Cooperative robotic teams are used in terrestrial applications and in the orbital and planetary environment [16]. Several applications involve cooperative robots where heterogeneous systems collaborate to achieve shared objectives [17,18]. Other HMRSs offer the option of expanding the entire mission scenario by (re-)configuring the systems involved with the help of so-called payload modules [19]. There is ongoing research on human–robot interaction involving multi-robot systems and astronauts on the Moon [20] and Mars [21].

Experiments demonstrate that the robots depicted in (Figure 2) show promising potential for planetary applications [22]. Figure 2a shows the several systems which are used to set up an In Situ Resource Utilization (ISRU) pilot plant. The goal was to work together to achieve common goals, including cooperative mission planning and executing actions for the transportation and assembly of the facility and its supporting infrastructure [23]. Figure 2b shows heterogeneous robotic systems in a field trial campaign in a desert in Utah. The cooperation of the systems was tested in a simulated mission scenario. The primary focus was on the execution of a semi-autonomous mission sequence (https://youtu.be/pvKIzldni68) (accessed on 28 March 2025) [19]. Figure 2c shows the team of cooperating autonomous robots during the field test in Lanzarote. The team explored hard-to-reach areas in Lanzarote to simulate planetary surfaces, such as lava tubes

on the Moon and mining tunnels on Earth (https://youtu.be/lEG1rQuOOI8) (accessed on 28 March 2025) [24].



Figure 2. Multi-robot teams—(**a**) Mantis (left) and Veles (right) [23] in Bremen; (**b**) SherpaTT in Utah; and (**c**) CoyoteIII (front) and SherpaTT (back) in Lanzarote. Image source: DFKI GmbH.

3.2. Contextual MDPs for Multi-Robot Systems

The use of context in the multi-agent setting remains relatively unexplored when compared to the bandit literature and single agent RL. The majority of work in this direction is investigating the theoretical aspects of including side information in the learning process, e.g., as contextual games [25,26]. Another strain of research tackles the more general problem of context-aware multi-agent systems (CA-MASs) [27]. Some experiments and analyses have also been performed over the application of context-aware methodologies to multi-agent RL scenarios such as trajectory generation in diverse multiple human–robot interaction scenarios [28] and the intelligent control of traffic lights [29].

To the best of the authors' knowledge, there exist no prior works that directly apply the formalism of CMDPs or contextual games to the problem of teaming in heterogeneous multi-agent systems.

4. Proposed Framework

The main objective of this work is to optimize planetary exploration missions by enhancing multi-robot teaming through an adaptive decision-making approach. The proposed framework addresses this challenge by applying contextual learning to the multi-agent scenario. Heterogeneity in physical capabilities enables robots to specialize in different types of tasks. In the best case scenario, a good mission planning algorithm will produce optimum results with little error. However, in the real-world case, uncertainties such as terrain variability, environmental changes, and hardware limitations can significantly impact performance. These challenges necessitate the incorporation of factors such as physical state, power consumption, status of sensors and communication, etc., into the decision-making framework. In essence, our method aims to accomplish the following goals:

1. Implement a decentralized task negotiation for a heterogeneous multi-robot team;
2. Incorporate real-time monitoring of hardware states (sensor noise, mechanical wear) into the task negotiation framework as a CMDP;
3. Improve robot and team adaptability;
4. Foster mutual support within robotic teams;
5. Optimize mission performance and team efficiency.

During autonomous missions, robotic systems operate independently, with minimal or no human intervention, either due to impracticality or by design. It is therefore important to monitor the hardware of the systems so that they are able to autonomously decide what to do in the event of a failure or other influences that endanger the systems. In the event of mechanical failures, e.g., wheels, cameras can check for visible damage. Appropriate sensors can provide on-board diagnostics for individual components installed inside the

robot. This includes monitoring battery capacity, sensor obstructions, communication loss, and malfunctions due to external influences like temperature and radiation.

4.1. Context-Aware Decision-Making Framework for Heterogeneous Multi-Robot Exploration

The proposed framework (Figure 3) enables an HMRS to collaboratively execute tasks in a dynamic extraterrestrial environment. Of the various hardware influences mentioned above, we consider input from the following four sources:

- **Robot State**: Robot operational status, pose and velocity, task status, and history.
- **Hardware (H/W) Parameters**: Sensor state, physical capabilities (affecting team composition, on-board diagnostics).
- **Environmental Dynamics**: Real-time factors like terrain type, Simultaneous Localization and Mapping (SLAM), weather conditions.
- **Communication State**: Status of connectivity between team members.

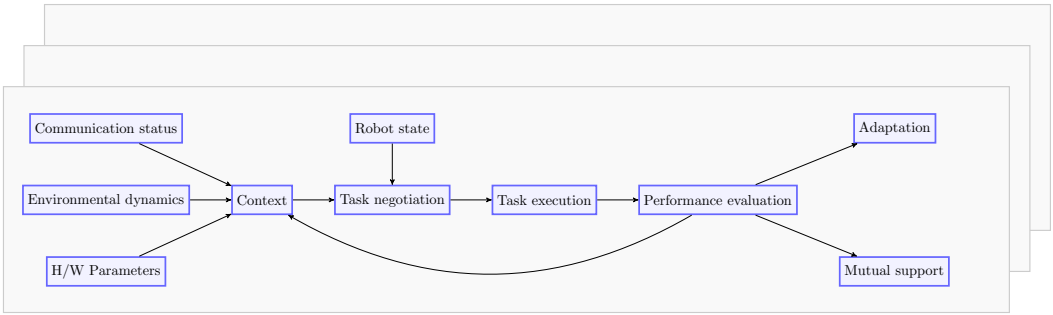


Figure 3. Context-aware decision-making framework for heterogeneous multi-robot teaming. A context is formalized based on the communication status, environmental dynamics, and robot hardware parameters as input. Then, the robot negotiates for tasks based on the current context and its high-level state. Executed tasks are evaluated to assess the adaptation and mutual support metrics. This performance evaluation further influences future decisions, thus closing the loop.

Based on these inputs, the context of the team and the states of an individual robot i at time t are determined. This state acts as the foundation for further decisions made by the robot within the given context and is dynamically updated as and when there is a change in input states. Once the context and state are established, the auction-based task negotiation algorithm then decides the respective robot's bid for the task(s) at hand. Therefore tasks are distributed among and executed by robots based on their contextual suitability.

4.2. Example Scenario

Consider a planetary exploration mission wherein a heterogeneous team of robots is deployed to explore a target area and transport payload modules between certain locations. The team consists of three robots, namely, SherpaTT, CoyoteIII, and Mantis [30]. CoyoteIII (Figure 2c—front) is a micro-rover which has advanced mobility capabilities in unstructured terrain. In this scenario, CoyoteIII acts as a scouting rover that maps large areas, creates terrain maps for safe navigation by other rovers, and identifies regions of interest for scientific investigation. SherpaTT (Figure 2b) is a walking–driving hybrid that can navigate in highly uneven terrain and transport payload modules from one place to another assisted by its 6-DOF arm. Mantis (Figure 2a) is a six-legged manipulation and locomotion system designed for a variety of complex tasks in difficult terrain. It can either function in locomotion mode, wherein it walks on all its six legs, or in manipulation mode by using one or both of its front legs as manipulators for handling modules or scientific samples.

Assuming the problem of planetary exploration with heterogeneous robot teams, we formalize the task as the following CMDP:

- Context space $\mathcal{C}$ consists of the team composition, each robot's hardware parameters, communication state, target area to explore, and weather conditions;
- State space $\mathcal{S}$ consists of high-level robot state of each robot team member;
- Action space $\mathcal{A}$ consists of high-level actions for each robot, e.g., target coordinates;
- Reward function $\mathcal{R}_c(s_t, a_t, s_{t+1})$ is the percentage of the target area that has been explored so far

Figure 4 shows how robot team composition can be represented as context. Using RL algorithms, e.g., Soft-Actor Critic (SAC) [31], we can now learn a context-dependent policy $\pi(a|s;c)$ to solve the tasks dependent on a given context c. By sampling two different sets of contexts $\mathcal{C}_{train}$ and $\mathcal{C}_{test}$, we can effectively and reliably evaluate how well the learnt policy can generalize to unseen situations. Additionally, the learnt context-dependent value-function can be used to inform the team selection process for the task at hand to select the ideal team composition for each new task.

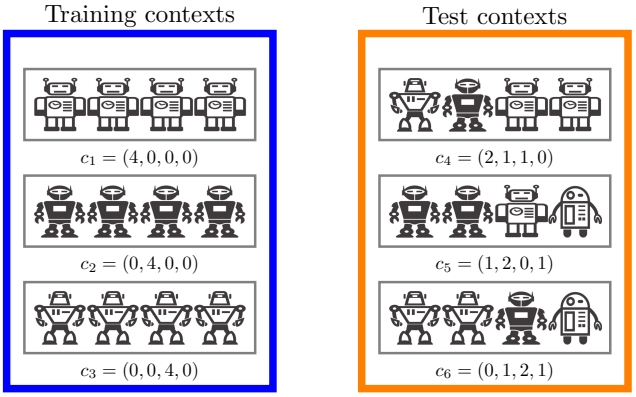


Figure 4. Team compositions are represented as context vectors indicating the number of robots for each robot type that is present in the team. Consider the situation where a team can consist of four robots of four different robot types. During training, the robot teams always consist of four equal robots; e.g., training context $c_1 = (4, 0, 0, 0)$ consists of four type-A robots. The goal is to learn a context-dependent policy that is able to solve a task with different team composition, potentially previously unseen ones. To evaluate the generalization to unfamiliar teams, a separate set of training sets can be defined in which the team may be heterogeneous, e.g., as $c_5 = (1, 2, 0, 1)$.

The performance of the robotic team and feedback loop will be discussed in the next section.

5. Evaluation

The team performance will be evaluated based on two factors in their behavior, namely, adaptation and mutual support.

5.1. Adaptation

Adaptation refers to a robot's ability to modify its behavior based on real-time feedback and the actions of other robots, measuring how effectively it responds to contextual changes to improve task efficiency. In the above scenario, CoyoteIII broadcasts a terrain map that helps the other two rovers to plan safe trajectories to their desired goal positions while avoiding unsuitable terrain like steep slopes. On the other hand, a dust storm could lead to camera distortion or dust accumulation on the solar panels, thereby increasing energy consumption and reducing energy storage capacity. In this case, the affected rovers switch to safer tasks like exploring flatter regions. Alternatively, an additional CoyoteIII can join the team to carry on the mapping task. Depending on task requirements, robots could enter or leave the team. Both cases lead to a change in team composition context. Identifying

and learning from this contextual change characterizes the adaptability of the robotic team. Adaptation can therefore be quantified by the percentage area explored, mission duration, damage incurred, and energy consumed.

5.2. Mutual Support

When a robot struggles to complete its assigned task due to an uncertainty induced by one or more of the input sources, another member of the team with the required resources can provide assistance, ensuring mission continuity. Mutual support can be accomplished by enabling robots to multi-task, autonomously repair, and reconfigure. For instance, CoyoteIII finds a lava tube that needs to be explored. Figure 2c depicts how SherpaTT tethers the CoyoteIII and gradually lowers it down the skylight. The mutual support factor also enhances overall team performance by dynamically redistributing physical resources among modular and reconfigurable robots, in which case the team composition context changes.

6. Conclusions and Outlook

This paper introduces a novel approach to multi-robot task negotiation, leveraging contextual reinforcement learning to enhance adaptability and task efficiency during planetary exploration. Our method aims to address the central research question of how incorporating team composition, hardware state, and environmental dynamics context into a contextual Markov Decision Process improves the resilience of heterogeneous multi-robot systems. The proposed framework is designed to maintain mission continuity, enabling the robotic team to overcome hardware faults and environmental challenges through real-time adaptation to situations and mutual support mechanisms.

Future work will involve an experimental evaluation of the framework, first in a simulated planetary environment and then at an analog test facility. A series of tests will assess task performance and coordination under various physical conditions based on the evaluation criteria mentioned in this paper. Context-aware decision-making for multi-agent systems is promising because it combines the advantages of heterogeneous teams with a contextual learning model.

Author Contributions: Introduction, A.S., M.L., and W.B.; background, M.L.; state of the art, A.S., M.L., and W.B.; proposed method, A.S., M.L., L.C.D., and W.B.; evaluation, A.S.; conclusion and outlook, A.S., M.L., and W.B.; supervision, F.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the Federal State of Bremen and the University of Bremen as part of the Humans on Mars Initiative by the German Ministry of Education and Research (BMBF) under grant 16KISK016 (Open6GHub) and by the German Research Foundation (DFG) under grant 500260669 (SCIL).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Authors Amrita Suresh and Melvin Laux are employed by the University of Bremen. Authors Wiebke Brinkmann and Leon C. Danter are employed by the company German Research Center for Artificial Intelligence (DFKI) GmbH. Frank Kirchner is affiliated with both, the DFKI and the University of Bremen. The remaining authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Gao, Y.; Chien, S. Review on space robotics: Toward top-level science through space exploration. *Sci. Robot.* **2017**, *2*, eaan5074. [CrossRef] [PubMed]
2. Balaram, J.; Aung, M.; Golombek, M.P. The ingenuity helicopter on the perseverance rover. *Space Sci. Rev.* **2021**, *217*, 56. [CrossRef]
3. Rizk, Y.; Awad, M.; Tunstel, E.W. Cooperative heterogeneous multi-robot systems: A survey. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–31. [CrossRef]
4. Seenu, N.; RM, K.C.; Ramya, M.; Janardhanan, M.N. Review on state-of-the-art dynamic task allocation strategies for multiple-robot systems. *Ind. Robot. Int. J. Robot. Res. Appl.* **2020**, *47*, 929–942.
5. Puterman, M.L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*; Wiley Series in Probability and Statistics; Wiley: Hoboken, NJ, USA, 1994. [CrossRef]
6. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; A Bradford Book: Cambridge, MA, USA, 2018.
7. Silver, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; et al. A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go through Self-Play. *Science* **2018**, *362*, 1140–1144. [CrossRef] [PubMed]
8. Vinyals, O.; Babuschkin, I.; Czarnecki, W.M.; Mathieu, M.; Dudzik, A.; Chung, J.; Choi, D.H.; Powell, R.; Ewalds, T.; Georgiev, P.; et al. Grandmaster Level in StarCraft II Using Multi-Agent Reinforcement Learning. *Nature* **2019**, *575*, 350–354. [CrossRef] [PubMed]
9. Hallak, A.; Di Castro, D.; Mannor, S. Contextual Markov Decision Processes. *arXiv* **2015**, arXiv:1502.02259.
10. Modi, A.; Jiang, N.; Singh, S.; Tewari, A. Markov Decision Processes with Continuous Side Information. In Proceedings of the Algorithmic Learning Theory, ALT 2018, Lanzarote, Canary Islands, Spain, 7–9 April 2018; Volume 83, pp. 597–618.
11. Kirk, R.; Zhang, A.; Grefenstette, E.; Rocktäschel, T. A Survey of Zero-shot Generalisation in Deep Reinforcement Learning. *J. Artif. Intell. Res.* **2023**, *76*, 201–264. [CrossRef]
12. Benjamins, C.; Eimer, T.; Schubert, F.; Mohan, A.; Döhler, S.; Biedenkapp, A.; Rosenhahn, B.; Hutter, F.; Lindauer, M. Contextualize Me—The Case for Context in Reinforcement Learning. *Trans. Mach. Learn. Res.* **2023** .
13. Otte, M.; Kuhlman, M.J.; Sofge, D. Auctions for multi-robot task allocation in communication limited environments. *Auton. Robots* **2020**, *44*, 547–584. [CrossRef]
14. Abdelnabi, S.; Gomaa, A.; Sivaprasad, S.; Schönherr, L.; Fritz, M. LLM-Deliberation: Evaluating LLMs with Interactive Multi-Agent Negotiation Games. 2023. Available online: https://dblp.org/rec/journals/corr/abs-2309-17234.html?view=bibtex (accessed on 4 December 2024).
15. Schillinger, P.; García, S.; Makris, A.; Roditakis, K.; Logothetis, M.; Alevizos, K.; Ren, W.; Tajvar, P.; Pelliccione, P.; Argyros, A.; et al. Adaptive heterogeneous multi-robot collaboration from formal task specifications. *Robot. Auton. Syst.* **2021**, *145*, 103866. [CrossRef]
16. Leitner, J. Multi-robot cooperation in space: A survey. In Proceedings of the 2009 Advanced Technologies for Enhanced Quality of Life, Iasi, Romania, 22–26 July 2009; pp. 144–151.
17. de la Croix, J.P.; Rossi, F.; Brockers, R.; Aguilar, D.; Albee, K.; Boroson, E.; Cauligi, A.; Delaune, J.; Hewitt, R.; Kogan, D.; et al. Multi-Agent Autonomy for Space Exploration on the CADRE Lunar Technology Demonstration. In Proceedings of the 2024 IEEE Aerospace Conference, Big Sky, MT, USA, 2–9 March 2024; pp. 1–14.
18. Burkhard, L.; Sakagami, R.; Lakatos, K.; Gmeiner, H.; Lehner, P.; Reill, J.; Müller, M.G.; Durner, M.; Wedler, A. Collaborative Multi-Rover Crater Exploration: Concept and Results from the ARCHES Analog Mission. In Proceedings of the 2024 IEEE Aerospace Conference, Big Sky, MT, USA, 2–9 March 2024; pp. 1–14.
19. Sonsalla, R.; Cordes, F.; Christensen, L.; Roehr, T.M.; Stark, T.; Planthaber, S.; Maurus, M.; Mallwitz, M.; Kirchner, E.A. Field testing of a cooperative multi-robot sample return mission in mars analogue environment. In Proceedings of the 14th Symposium on Advanced Space Technologies in Robotics and Automation (ASTRA), Leiden, The Netherlands, 20–22 June 2017.
20. Imhof, B.; Hoheneder, W.; Ransom, S.; Waclavicek, R.; Davenport, B.; Weiss, P.; Gardette, B.; Taillebot, V.; Gobert, T.; Urbina, D.; et al. Moonwalk-human robot collaboration mission scenarios and simulations. In Proceedings of the AIAA SPACE 2015 Conference and Exposition, Pasadena, CA, USA, 31 August–2 September 2015; p. 4531.
21. Suresh, A.; Beck, E.; Dekorsy, A.; Rückert, P.; Tracht, K. Human-Integrated Multi-Agent Exploration Using Semantic Communication and Extended Reality Simulation. In Proceedings of the 2024 10th International Conference on Automation, Robotics and Applications (ICARA), Athens, Greece, 22–24 February 2024; pp. 419–426.
22. Brinkmann, W.; Danter, L.; Suresh, A.; Yüksel, M.; Meder, M.; Kirchner, F. Development strategies for multi-robot teams in context of planetary exploration. In Proceedings of the 2024 International Conference on Space Robotics (iSpaRo), Luxembourg, 24–27 June 2024; pp. 64–69.

23. Govindaraj, S.; Brinkmann, W.; Colmenero, F.; Nieto, I.S.; But, A.; de Benedetti, M.; Danter, L.; Alonso, M.; Heredia, E.; Lacroix, S.; et al. Building a lunar infrastructure with the help of a heterogeneous (semi) autonomous multi-robot-team. In Proceedings of the 72nd International Astronautical Congress (IAC-2021), Dubai, United Arab Emirates, 25–29 October 2021.
24. Vögele, T.; Ocón, J.; Germa, T.; Govindaraj, S.; Perez-Del-Pulgar, C.J.; Haugli, F.B.; Domínguez, R.; Danter, L.C.; Babel, J.; Dragomir, I.; et al. COROB-X: Demonstration of a cooperative robot team in extensive field tests. In Proceedings of the ASTRA 2023, Leiden, The Netherlands, 18–20 October 2023; pp. 1–9.
25. Sessa, P.G.; Bogunovic, I.; Krause, A.; Kamgarpour, M. Contextual Games: Multi-Agent Learning with Side Information. In Proceedings of the Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, Virtual, 6–12 December 2020.
26. Maddux, A.M.; Kamgarpour, M. Multi-Agent Learning in Contextual Games under Unknown Constraints. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Palau de Congressos, Valencia, Spain, 2–4 May 2024; Volume 238, pp. 3142–3150.
27. Du, H.; Thudumu, S.; Vasa, R.; Mouzakis, K. A Survey on Context-Aware Multi-Agent Systems: Techniques, Challenges and Future Directions. *arXiv* **2024**, arXiv:2402.01968.
28. Xu, Z.; Zhou, R.; Yin, Y.; Gao, H.; Tomizuka, M.; Li, J. MATRIX: Multi-Agent Trajectory Generation with Diverse Contexts. In Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA), Yokohama, Japan, 13–17 May 2024; pp. 12650–12657. [CrossRef]
29. Zhu, R.; Wu, S.; Li, L.; Lv, P.; Xu, M. Context-Aware Multiagent Broad Reinforcement Learning for Mixed Pedestrian-Vehicle Adaptive Traffic Light Control. *IEEE Internet Things J.* **2022**, *9*, 19694–19705. [CrossRef]
30. Brinkmann, W.; Bartsch, S.; Sonsalla, R.U.; Cordes, F.; Kuehn, D.; Kirchner, F. Advanced Robotic Systems in the Context of Future Space Exploration. In Proceedings of the 69th International Astronautical Congress 2018, (IAC-2018), Bremen, Germany, 1–5 October 2018.
31. Haarnoja, T.; Zhou, A.; Abbeel, P.; Levine, S. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; Volume 80, pp. 1861–1870.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.