# Towards Reducing Human Supervision in Fielded Multi-Agent Systems

Harel Biggie, and Christoffer Heckman*

***Abstract*— Recent advancements in perception, control, and planning have enabled single agents to partially explore unknown complex environments such as caves, and mines. In this work we exam the performance of a heterogeneous multi-agent navigation stack developed by Team MARBLE from CU Boulder as part of the DARPA Subterranean Challenge. The heterogeneous team consisted of two wheeled Clearpath Husky A200 vehicles and two quadraped Boston Dynamics Spot platforms. Coordination between the robots was performed using a mixture of autonomy and limited human supervision. We analyze performance of the system at the final competition for the DARPA SubT challenge and specifically look at areas where autonomous coordination fell short and human intervention was required. From this analysis we suggest future directions for how to shift towards a fully autonomous solutions for decision making in heterogeneous systems using contextual and semantic information.**

Fig. 1: Huksy A2000 (right) and Spot (left) configured for multi-agent exploration at the SubT Final event.

## I. Introduction

Recent advancements in planning, control, and perception have enabled robotic autonomy stacks capable of navigating both structured and unstructured environments. The DARPA Subterranean challenge (SubT) [1] was a three year long effort designed to spark innovation in the technical areas of autonomy, perception, networking and mobility for mobile robot exploration [2]. The challenge was modeled within the context of search and rescue and teams of robots needed to explore unknown underground environments while searching for objects that would indicate human presence. Specifically, the primary objective of the mission was "looking" which earned a point after each object was correctly identified and a secondary objective "exploring" was to produce a 3D volumetric map of the environment.

Under the "mobility" technical area, DARPA required robots to traverse challenging obstacles such as constrained passages, vertical shafts, steps, and mud. The variety of these obstacles combined with the need to build out communication infrastructure for transmitting data both externally and between robots necessitated the use of heterogeneous robot teams. No single platform had both the terrain capabilities and payload required to perform the entire mission.

Effective coordination amongst different robots is paramount to efficiently exploring previously unknown areas and completing the "looking" and "exploration" objectives of the mission. Coordination was made especially difficult in the SubT challenge due to a lack of a priori knowledge about the environment, limited communication bandwidth and complex environmental topologies. Robots were primarily required to explore the environment fully autonomously with the exception of a single human supervisor who could issue alternative commands when communications were available. In this work we take a look at the different multi-agent decision making approaches used by teams at the SubT challenge. Specifically, we take a closer look at the effectiveness of Team MARBLE's solution.

Team MARBLE placed third at the final competition with 18 out of 40 artifacts located and 48% of the environment explored. During the course of the 60 minute run the the human supervisor only intervened five times across the fleet of four robots. The rest of the time the robots managed themselves using an internal decision making framework. From our analysis we find:

1) The mission management system employed by Team MARBLE enabled autonomous exploration in complex subterranean environments with minimal human supervision.
2) Two critical interventions which lasted for 14 minutes of the 60 minute prize run led to a significant increase in exploration area. Despite the relatively limited amount of human intervention, such intervention is still required for peak system performance.
3) Future multi-agent mission management systems need to be able to change system behaviors based on risk assessments within the objectives of the mission.
4) Semantic information can play an instrumental role in assessing a situation and this type of information should be incorporated in future multi-agent coordination strategies.

This work was supported through the DARPA Subterranean Challenge, cooperative agreement number HR0011-18-2-0043. and by the National Science Foundation under Grant No. 1764092.

*All authors are with the Autonomous Robotics and Perception Group at the University of Colorado Boulder, Boulder Colorado, USA. Corresponding author: christoffer.heckman@colorado.edu

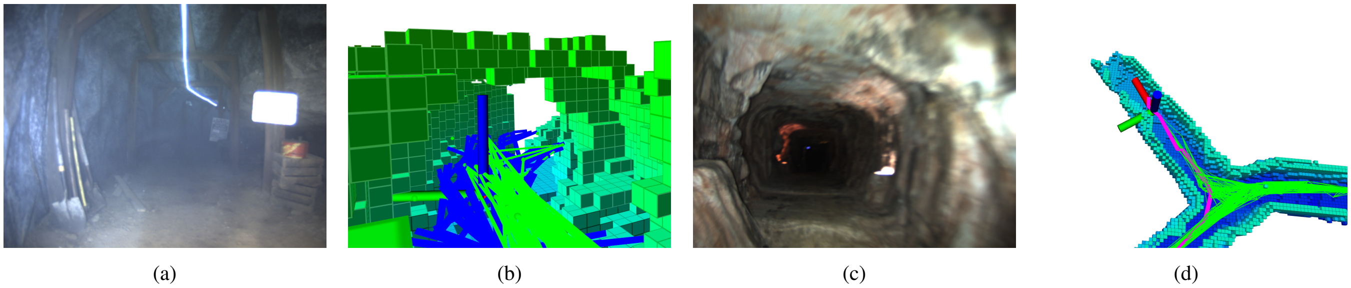

Fig. 2: Camera views of the foggy area (a) and the narrow passage (c). The maps for each area are shown in (b) and (d) respectively. The green lines show the planning tree and the blue lines show perspective paths that are being evaluated.

Finally, we discuss potential research directions to overcome these limitations and alleviate the need for human supervision.

## II. RELATED WORKS

Typical key elements that define a mulit-agent robot system (MARS) are the type of agent, the control architecture, and the communication strategy [3]. Types of agents are broadly categorized into homogeneous and heterogeneous while control architectures are categorized into decentralized, centralized, and hybrid which is a mixture of both centralized and decentralized control. Communication can either be implicit where interactions are made via sensing, or explicit where information is directly transferred between agents. The SubT challenge largely relied on heterogeneous fleets of robots using hybrid control architectures with explicit communication. As such we will focus our review on these types of systems.

A comprehensive survey on the state of MARS systems from 2018 can be found in [3] where the authors make recommendations for future work of which two are directly addressed by the SubT challenge. First, the authors recommend that suitable communication protocols and network control systems need to be implemented to avoid time delays during transmission of information among agents. A second recommendation was that more work needs to be done at the intersection of implicit and explicit control.

### *A. Approaches to Multi-Agent Coordination at the SubT Challenge*

Traditionally, single agent autonomous systems are controlled using finite-state machines (FSMs) [4] which select teams at the SubT challenge made use of [5], [6]. Team Cerberus initially augmented FSMs to include transitions and states based on inputs from other agents [7] for a form of explicit coordination.

Another mission management solution comes in the form of behavior trees which are a powerful way of modeling tasks with clear explanations as to why each task is chosen due to the tree like structure [8], [9]. For these reasons a variety of teams used behavior trees to model their multi-decision making pipelines [10], [7], [11], [7]. Examples of shared data that could influence the robot's behavior include frontier points for further exploration [12], volumetric maps [10], and reports of artifacts present in the environment.

Beyond decision making mechanisms these multi-agent systems needed the ability to efficiently select and transmit information inside of bandwidth constrained environments. The majority of teams used the Robot Operating System (ROS) as a backbone operating system for their autonomy stacks which by default uses TCP networking for single agent message passing. This approach does not scale well in connectivity limited environments and teams either adapted other message passing mechanisms [13], [14] or developed their own [15]. These architectures directly address the networking challenges identified in [3].

Because the challenge allowed for a single human in the loop, many of these systems were able to perform different behaviors based on input states given by the human. Team CoSTAR developed a "copilot" system to aid their human supervisor in controlling the full fleet of robots. This system provided a graphical user interface, task scheduling and ten operating modes for the system [16], [17]. Generally, teams constructed a human supervisor interface and were able to send waypoints to their robots or take limited remote control when bandwidth allowed [7].

## III. SYSTEM OVERVIEW

Team MARBLE focused on an autonomy first paradigm for mission operations in which we assumed communication would be intermittent at best and therefore the system needed to operate with as little human intervention as possible. In the context of the DARPA challenge, the team of robots was deployed into the final course for a total of 60 minutes. During the course of the run, only a single human supervisor was able to monitor the team of robots and issue high level commands when communications were available. In this section we provide a high level system overview of the key components relevant to multi-agent coordination and refer the reader to [18], [19] for a detailed description of the entire system.

*Platform Roles:* Team Marble utilized two Boston Dynamics Spot quadruped robots and two Clearpath Husky A200 wheeled robots with example configurations shown in

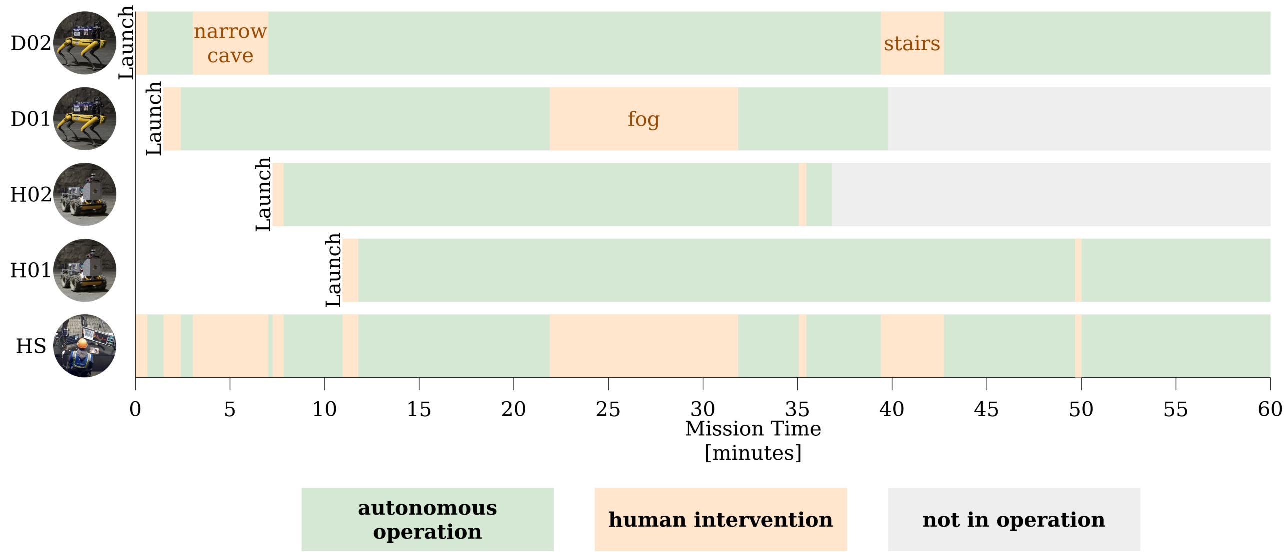


Fig. 3: Mission timeline for the 60 minute final prize run at the SubT challenge. HS represents the Human Supervisor and key mission interventions (stairs and fog) are labelled.

Figure 1. The Spot platforms were used primarily for both "looking" and "exploring" due to their high ground clearance and ability to climb stairs. Meanwhile the Husky platforms were equipped with similar exploration capabilities and they also carried six communication beacons for building out a communication infrastructure. Spot platforms were deployed first in a "breadth" first fashion and were followed by the Huskies which built out communication infrastructure and subsequently provided a more in depth search of explored areas.

*Mapping and Planning:* Volumetric maps were based off on the Octomap [20] package with modifications to support difference based map transmission [21] and semantic labeling [15]. Specifically, semantic labels were added for staircases and rough terrain which were used by the RRT graph based planner. The sampling based planner [22] prioritizes going to areas which provide new volumetric gain for the map using a frontier based approach [12]. Notably, the planner was tuned conservatively by adjusting the robot radius for collision checks and the maximum step height the robots would traverse in order to mitigate risk to the agent. The planner also accepts pose graphs from other agents in order to direct the current agent to areas which have not been previously explored.

*Communications:* Robots used a deployed mesh network to both communicate amongst themselves and to transmit data back to the human supervisor. Key features of the network included rapid reconnection times in dynamic environments, and data prioritization and the full details can be found in [15]. Transmitted data was controlled by the Multi-Agent Data Collaboration for Autonomous Teams (MADCAT) framework as described in [18]. MADCAT aggregates goal points, artifact reports, and maps from all agents and uses weighted time based deconfliction to prioritize messages.

*Mission Management:* Individual agents take the transmitted elements from MADCAT and use a series of weighted mission management approach called Behaviors, Objectives and Binary states for Coordination of Autonomous Tasks (BOBCAT) [23] to make decisions. Decisions are made based on the weights of different system monitors such as the number of detected artifacts, or the time since the last artifact was reported. The objective (explore, report an artifact, go home, extend communications) with the highest set of weights is selected and the corresponding behavior is executed. The framework is also responsible for selecting when to utilize the goals from the default exploration mode of the planner or a human supervisor provided goal which is transmitted via MADCAT. The system also enabled the human to teleoperate a robot in high bandwidth situations using an Xbox controller.

## IV. RESULTS

We present the results from Team MARBLE's competition run at final event for the SubT. As seen in Figure 4 the final event course was divided into three distinct sections: tunnel, urban and cave. The figure depicts a simulated run of the finals course inside of the DARPA SubT simulator [24] as well as the actual explored area from the final run. We note, that the simulated run did not have any human supervision and the robots relied solely on the built in autonomy stack. From this figure it is clear that the exploration capabilities of the simulated run were similar to the real world inside of the urban environment. However, in both the tunnel and the cave sections of the course the real-world system explored substantially more area than the simulated system.

Figure 3 contains a timeline illustrating where the system operated with and without the input of the human agent. All robot launches were conducted manually to ensure the robots were fully functional before entering the course. For the purpose of this discussion we will not consider the launch

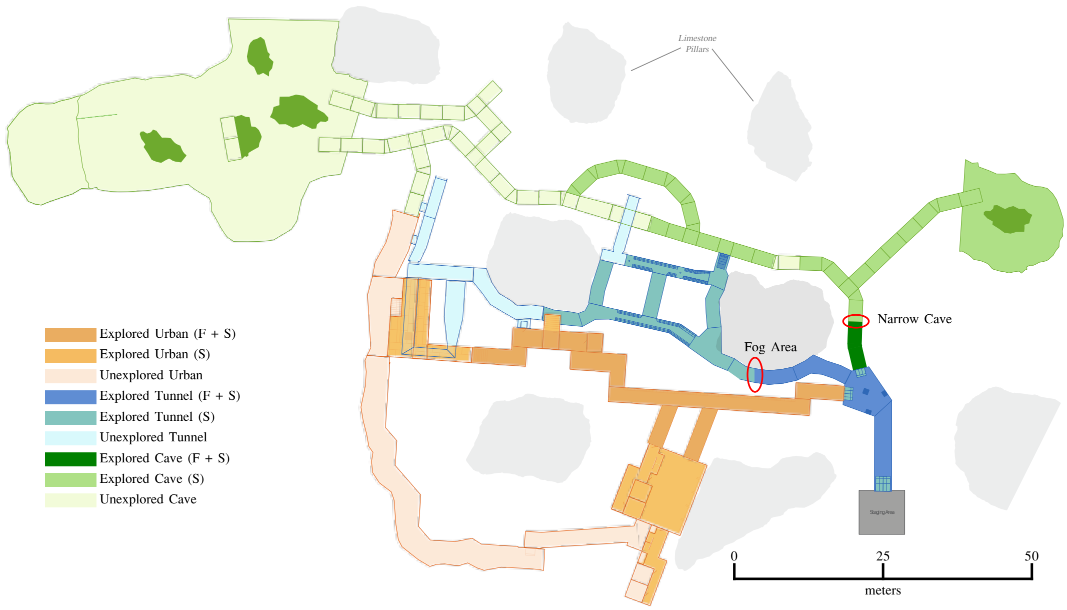


Fig. 4: Map of the course at the DARPA SubT final event divided into tunnel, urban and cave sections. A comparison of the explored area between a simulated (S) run with no human intervention and Team MARBLE's final event run (F) is shown.

sequence as an intervention which results in a total of 5 interventions across the 60 minute run. A summary of all interventions can be found in Table I. The first intervention occurred approximately 3 minutes into the run where the human supervisor manually navigated one of the Spots (D02) through a narrow cave shown in Figure 2c. The second major intervention was at 22 minutes where the human supervisor teleoperated the other Spot robot (D01) traverse through a foggy area shown in Figure 2a. Both of these areas are outlined in red on the full course map shown in Figure 4.

Interventions 3 and 5 both involved a robot planning towards the course entrance (also the exit) and were a result of a too small exclusion zone in the planning parameters. These interventions did not significantly impact the overall explored area. Intervention 4 was a manual attempt to guide the D02 down a staircase made by the human supervisor. The stair detection capabilities of the system were limited to upwards facing staircases as described in [15] which prevented the robot from autonomously traversing these stairs.

| Intervention | Agent | Duration | Goal |
|---|---|---|---|
| 1 | D02 | 240s | Enter narrow cave corridor |
| 2 | D01 | 598s | Enter foggy tunnel area |
| 3 | H02 | 24s | Avoid course exit |
| 4 | D02 | 200s | Walk down stairs |
| 5 | H01 | 21s | Avoid course exit |

TABLE I: List of the five robot interventions executed by the human supervisor during the final event prize round. All of the interventions except for walking down the stairs were successful. Beyond these interventions the robots were actively being managed by the onboard mission management system (BOBCAT).

Figure 2d shows the volumetric gain of the aggregate map from all robots with respect to the mission time. Comparing the gains at approximately 11 and 25 to the timeline in Figure 3 we can see that the volumetric gain significantly increased after the "narrow cave" and "fog" intervention.

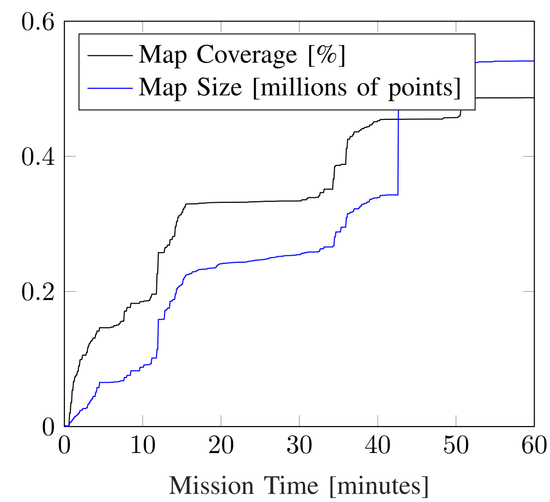


Fig. 5: Volumetric gain of all four robots (D01,D02,H01,H02) with respect to time. The map size is shown in blue and the coverage percentage is shown in black. The coverage percentage is defined by the fraction of points within 1m of a ground truth point.

## V. DISCUSSION

Based on the key highlighted "fog" and "narrow cave" areas of Figure 4 we can see that those interventions contributed to a significant increase in the exploration area of the system. In this section we explore the limitations in the mission management system that prevented the robots from making the decision to traverse these areas on their own.

Taking a closer look at the "narrow cave" we can see from Figure 2c the robot did not extend its planning graph (shown in green) into the area of the narrow cave. The reason for this was a conservative value for the robot radius, which mandates the minimum distance threshold the robot needs to be away from an object. This parameter was chosen before the mission in order to minimize the risk of the platform falling over. In contrast, the human was able to assess the risks and potential benefits of the situation and make the decision to manually send the robot through. To achieve a similar level of performance a mission management system would have needed to instruct the planner to temporarily reduce the robot radius which would have enabled an automated plan through the area. Despite a seemingly trivial planning solution, making the initial decision is nearly impossible given the information available to the BOBCAT system.

The human was able to make the decision due to the context of the map. The fact that the "narrow cave" occurred near the beginning of the corridor suggests that more area exists behind it. Additionally, cave environments typically have narrow passages that open into wider areas. Providing this level of semantic information would have assisted in making this decision autonomously. An existing body of research provides the mechanisms to build these types of rich semantic maps [25]. However, more fundamentally, behavior trees, state machines, and weighted objective based mission managers like BOBCAT do not have the mechanisms to reason over rich semantic information. There are simply too many combinations of scenarios dictated by information such as topology and environment classification to predetermine the possible actions a robot should take.

Similarly in the fog instance, semantic information could have been used to adjust the planning parameters. Fundamentally, as seen from Figure 2b, the planner was able to propagate its tree through the foggy area but did not direct the robot into that area due to low volumetric gain. Instead, the human supervisor manually drove the robot through the area and as evident by the continued exploration after this intervention the onboard mapping, state estimation and navigational systems were robust to the limited perceptual abilities inside of the fog. This intervention took place approximately 22 minutes into the mission at a point where the map coverage had plateaued (Figure 5. At this point, D01 had limited additional areas to explore without going through the fog and as such the human supervisor determined it was worthwhile to risk the stability of the onboard perception system. Clearly, the risk payed off here but even if it hadn't, the other agents of the system would have continued exploring and the loss of a single agent would not have been mission ending. To enable traversal through fog a mission management system would have been required here to override the initial volumetric gain parameters of the planner. Making this decision entails understanding the risks fog poses to the perception suite as well as understanding that within the context of the rest of the mission and the other agents, there was no better task for D01.

In both of these instances, nothing in the core autonomy stack of the robot (planning and localization) prevented the robot from having the theoretical capability of traversing these areas on its own. However, the performance of the planner is artificially limited by tunable parameters such as volumetric gain, and the robot radius. These parameters were chosen prior to the run to minimize risk and "overriding" them was left up to the human supervisor. In order to remove the dependency on human input multi-agent frameworks such as BOBCAT need to be able to evaluate risk based on semantic information and the context of the mission.

## VI. Conclusions and Future Research Directions

We have shown that that within the context of search and rescue in unknown environments, human supervision is still critical for mission success. Additionally, we highlight the need for better risk based mission management systems that are able to both account for the context of the mission and semantic information. The sheer number of potential scenarios suggests traditional hand engineered mission management systems are insufficient. A possible way of addressing this challenge would be to leverage deep reinforcement learning techniques (DRL) for exploration [26], [27]. More recently DRL has been been shown to outperform traditional exploration methods in cluttered environments within the context of search and rescue [28]. While still largely limited to simulation or environments that contain relatively simple terrain challenges, these methods show incredible promise. To overcome the limitations of current mission management systems we need to explore innovative ways of leveraging the efficient exploration learned by methods such as [28] and combining them with more traditional planners that are already tested in harsh complex environments.

## References


[1] V. L. Orekhov and T. H. Chung, "The darpa subterranean challenge: A synopsis of the circuits stage," *Field Robotics*, pp. 735–747, 2022.

[2] DARPA, "DARPA Subterranean Challenge," 2022. https://www.darpa.mil/program/darpa-subterranean-challenge.

[3] Z. H. Ismail and N. Sariff, "A survey and analysis of cooperative multi-agent robot systems: challenges and directions," in *Applications of Mobile Robots*, pp. 8–14, IntechOpen, 2018.

[4] A. Kurt and Ü. Özgüner, "Hierarchical finite state machines for autonomous mobile systems," *Control Engineering Practice*, vol. 21, no. 2, pp. 184–194, 2013.

[5] Y.-W. Huang, C.-L. Lu, K.-L. Chen, P.-S. Ser, J.-T. Huang, Y.-C. Shen, P.-W. Chen, P.-K. Chang, S.-C. Lee, and H.-C. Wang, "Duckiefloat: a collision-tolerant resource-constrained blimp for long-term autonomy in subterranean environments," *arXiv preprint arXiv:1910.14275*, 2019.

[6] I. D. Miller, F. Cladera, A. Cowley, S. S. Shivakumar, E. S. Lee, L. Jarin-Lipschitz, A. Bhat, N. Rodrigues, A. Zhou, A. Cohen, *et al.*, "Mine tunnel exploration using multiple quadrupedal robots," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2840–2847, 2020.

[7] M. Tranzatto, F. Mascarich, L. Bernreiter, C. Godinho, M. Camurri, S. Khattak, T. Dang, V. Reijgwart, J. Loeje, D. Wisth, *et al.*, "Cerberus: Autonomous legged and aerial robotic exploration in the tunnel and urban circuits of the darpa subterranean challenge," *arXiv preprint arXiv:2201.07067*, 2022.

[8] M. Colledanchise and P. Ögren, *Behavior trees in robotics and AI: An introduction*. CRC Press, 2018.

[9] M. Colledanchise, D. Almeida, and P. Ögren, "Towards blended reactive planning and acting using behavior trees," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 8839–8845, IEEE, 2019.

[10] S. Scherer, V. Agrawal, G. Best, C. Cao, K. Cujic, R. Darnley, R. DeBortoli, E. Dexheimer, B. Drozd, R. Garg, *et al.*, "Resilient and modular subterranean exploration with a team of roving and flying robots," *Submitted to the Journal of Field Robotics*, vol. 2, no. 3, p. 6, 2021.

[11] K. Otsu, S. Tepsuporn, R. Thakker, T. S. Vaquero, J. A. Edlund, W. Walsh, G. Miles, T. Heywood, M. T. Wolf, and A.-A. Agha-Mohammadi, "Supervised autonomy for communication-degraded subterranean exploration by a robot team," in *2020 IEEE Aerospace Conference*, pp. 1–9, IEEE, 2020.

[12] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97.'Towards New Computational Principles for Robotics and Automation'*, pp. 146–151, IEEE, 1997.

[13] M. Schwarz, M. Beul, D. Droeschel, S. Schüller, A. S. Periyasamy, C. Lenz, M. Schreiber, and S. Behnke, "Supervised autonomy for exploration and mobile manipulation in rough terrain with a centaur-like robot," *Frontiers in Robotics and AI*, vol. 3, p. 57, 2016.

[14] D. Tardioli, R. Parasuraman, and P. Ögren, "Pound : A multi-master ROS node for reducing delay and jitter in wireless multi-robot networks," *Robotics and Autonomous Systems*, vol. 111, pp. 73–87, 2019. QC 20181210.

[15] H. Biggie and S. McGuire, "Heterogeneous ground-air autonomous vehicle networking in austere environments: Practical implementation of a mesh network in the darpa subterranean challenge," *arXiv preprint arXiv:2203.12832*, 2022.

[16] A. Agha, K. Otsu, B. Morrell, D. D. Fan, R. Thakker, A. Santamaria-Navarro, S.-K. Kim, A. Bouman, X. Lei, J. Edlund, *et al.*, "Nebula: Quest for robotic autonomy in challenging environments; team costar at the darpa subterranean challenge," *arXiv preprint arXiv:2103.11470*, 2021.

[17] M. Kaufmann, T. S. Vaquero, G. J. Correa, K. Otstr, M. F. Ginting, G. Beltrame, and A.-A. Agha-Mohammadi, "Copilot mike: An autonomous assistant for multi-robot operations in cave exploration," in *2021 IEEE Aerospace Conference (50100)*, pp. 1–9, IEEE, 2021.

[18] M. T. Ohradzansky, E. R. Rush, D. G. Riley, A. B. Mills, S. Ahmad, S. McGuire, H. Biggie, K. Harlow, M. J. Miles, E. W. Frew, *et al.*, "Multi-agent autonomy: Advancements and challenges in subterranean exploration," *Field Robotics*, pp. 1068–1104, 2022.

[19] H. Biggie, E. R. Rush, D. G. Riley, S. Ahmad, S. McGuire, H. Biggie, K. Harlow, M. J. Miles, E. W. Frew, *et al.*, "Flexible supervised autonomy for exploration in subterranean environments," *Field Robotics*, 2022.

[20] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: An efficient probabilistic 3d mapping framework based on octrees," *Autonomous robots*, vol. 34, no. 3, pp. 189–206, 2013.

[21] D. G. Riley, "MARBLE Mapping."

[22] S. Ahmad and J. S. Humbert, "Efficient sampling-based planning for subterranean exploration," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2022.

[23] D. G. R. Ii and E. W. Frew, "Assessment of coordinated heterogeneous exploration of complex environments," in *2021 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 138–143, IEEE, 2021.

[24] F. Open Robotics, "SubT Simulator."

[25] I. Kostavelis and A. Gasteratos, "Semantic mapping for mobile robotics tasks: A survey," *Robotics and Autonomous Systems*, vol. 66, pp. 86–103, 2015.

[26] J. Zhang, L. Tai, M. Liu, J. Boedecker, and W. Burgard, "Neural slam: Learning to explore with external memory," *arXiv preprint arXiv:1706.09520*, 2017.

[27] L. Tai and M. Liu, "A robot exploration strategy based on q-learning network," in *2016 ieee international conference on real-time computing and robotics (rcar)*, pp. 57–62, IEEE, 2016.

[28] F. Niroui, K. Zhang, Z. Kashino, and G. Nejat, "Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 610–617, 2019.