Bio-inspired Dynamic and Decentralized Online Learning in Uninformed Heterogeneous Multi-Agent Environments

Angel Sylvester
University of Minnesota
sylve057@umn.edu

Abstract

Dynamic adaptation and learning, akin to natural organisms, is crucial for robots operating in real-world scenarios like search and rescue missions. We propose a solution combining intuition from embodied evolution and Bayes theory to promote flexible exploration in foraging tasks. Our investigation focuses on three main areas: 1) leveraging communication and prior knowledge to develop adaptable strategies in agent groups, 2) addressing challenges from sparse rewards or limited data availability, and 3) developing methods for concurrent evaluation and training in a single iteration, filling a current gap in learning-based solutions. Future directions include exploring decentralized coordination among agents and incorporating assistance based on prospective memory and altruism in multi-agent reinforcement learning.

1 Introduction

Swarm robotics draws inspiration from biological entities like bees, beetles, and foraging ants. Each robot is equipped with a controller that guides its actions based on environmental feedback, allowing it to perform tasks like collecting food and adapting to the surroundings. From a purely biological context, it has been shown that organisms change behaviors dynamically in response to incentives and punishments from the environment [Gordon, 2016; Pasquier and Grüter, 2016]. Over longer periods of time, this is manifested through evolutionary adaptations that promote survival by tailoring traits to suit specific environments [Gordon, 2020].

A frequently encountered and extensively studied challenge within swarm robotics involves the foraging scenario, wherein a collective behavior emerges as a group of robots search, gather, and transport food items in their environment. The overarching objectives for this task are enhanced search efficiency, adaptability, and scalability.

2 Related Work

Many foraging methods take cues from social insects to address the computational challenges posed by learning frameworks and the lack of ability to generalize, specifically in dynamically changing environments. Examples of algorithmic approaches inspired by social insects include the artificial bee colony algorithm, artificial bacteria foraging, ant colony optimization, and particle swarm optimization. These methods are not computationally intensive and can adapt online without prior knowledge, but do have drawbacks such as constant overhead or centralization of the computation. They are useful for warehouse situations where across scenarios collection locations are standardized.

In scenarios like search and rescue operations, however, where the system lacks prior knowledge or where object distribution is mostly random and environments are varied, solutions that steer towards efficient search behaviors can enhance collection effectiveness.

This work specifically delves into methods found in *embodied evolution* [Bredeche *et al.*, 2018]. This contrasts with traditional evolutionary robotics which employs evolution in a sequential centralized optimization paradigm. Embodied evolution, in contrast, is characterized by decentralization, online execution, and parallelization. The parallelization specifically allows multiple agents in a shared environment to perform actions and evolve concurrently.

3 Main Research Questions/Contributions

Even within the realm of *embodied evolution*, a clear distinction exists between training and evaluation phases. When prior knowledge of the environments is limited, there is value in learning during execution, allowing for mistakes to be made and learning to happen within a time-sensitive context. To do so, we dive deeper into existing work in biological insights on dynamic learning and adaptation, by integrating Bayesian updating and optimal foraging theory as an alternative to explicit forms of long-term memory [Sylvester and Gini, 2023].

We introduce ADAPT-GA (Adaptive Dynamic Agent Parameter Tuning Genetic Algorithm), which combines an online evolutionary algorithm which leverages local communication for parameter sharing at set intervals throughout the task, along with Bayesian-like intuition to bias search strategies for the current environment. Key parameters shared by the robots include speed, object reward, and exploration penalty, shaping strategy-based decision-making.

Robots select an exploration strategy based on probability distributions which get updated based on the current performance. The strategies available to each robot are: forward persistence, circular persistence which divides the space into four bearings to undergo counterclockwise/clockwise movement, and correlated random walk which moves randomly but with a bias for the current direction. The strategies are weighted depending on performance. Instantaneous updates to the weights bias toward better-performing strategies.

Our solution addresses the foraging problem by dynamically updating robot exploration behaviors in response to events that affect performance and using a genetic algorithm to adjust parameter values, as illustrated in Fig. 1.

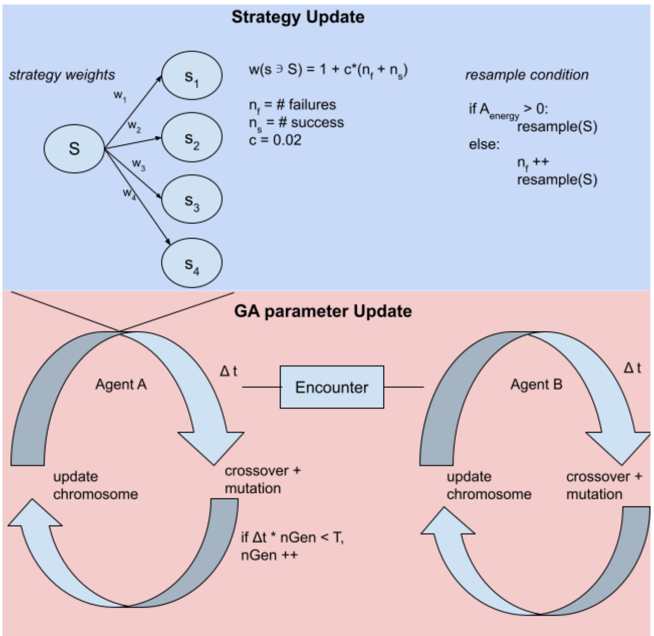

Figure 1: High-Level Overview of ADAPT-GA: The blue component represents strategy weights in a Bayes-based strategy distribution, with strategy changes triggered by an agent's energy status dropping below zero. The red component involves online updates of decision-making parameters and facilitates the exchange of performance information with nearby agents, potentially influencing the crossover phase for subsequent generations

Our key contributions are twofold: first, we emphasize the role of local communication in enhancing decision-making among individual agents, and we integrate a Bayesian-based intuition and online evolutionary algorithm to inform the behavior of individual robots based on their personal foraging performance and the decision-making parameters that are shared by other agents throughout the exploration task. Second, we seamlessly integrate evaluation into training, aiming to achieve a balance between exploration and exploitation in a single iteration even in dynamic environments.

In different environments like random dispersed objects, object clusters, and obstacle-laden urban settings, our algorithm, ADAPT-GA, demonstrates superior performance compared to other online-based bio-inspired methods used for foraging tasks [Sylvester and Gini, 2023]. Our ablation studies show that communication appears to enhance performance, further validating the importance of inter-agent interactions for decision-making. We have also shown that increased population density and potential for congestion do not negatively impact performance stability.

4 Future Directions

Our upcoming work involves expanding upon our current progress with ADAPT-GA. We are particularly interested in delving deeper into leveraging local communication to enhance decision-making among individuals, while also preserving decentralized learning for coordination strategies. Our approach will involve refining coordination strategies using Bayesian updating and incorporating drift (gamma decay). Furthermore, we aim to generate feasible trajectories aligned with these coordination strategies by adapting techniques from rapidly exploring decision trees or Monte Carlo decision trees, maintaining online decentralized learning while producing a sequence of actions. Additionally, we will focus on enhancing the agents' ability to anticipate future scenarios and adjust coordination strategies accordingly based on real-time environmental cues.

We are also intrigued by exploring the applicability of these insights to address social dilemma challenges within multi-agent reinforcement learning (MARL). By employing real-time environmental reasoning, we aim to dynamically shape rewards to foster altruistic actions aligned with the needs, statuses, or challenges faced by others. This approach highlights traits like restraint and prospective memory, aiming to implicitly promote such behaviors rather than enforce them through contractual obligations.

Ethical Statement

There are no ethical issues.

Acknowledgments

Support is gratefully acknowledged from the Minnesota Supercomputing Institute.

References

[Bredeche *et al.*, 2018] Nicolas Bredeche, Evert Haasdijk, and Abraham Prieto. Embodied evolution in collective robotics: A review. *Frontiers in Robotics and AI*, 5, 2018.

[Gordon, 2016] Deborah M. Gordon. The evolution of the algorithms for collective behavior. *Cell Systems*, 3(6):514–520, December 2016.

[Gordon, 2020] Deborah M Gordon. Movement, encounter rate, and collective behavior in ant colonies. *Annals of the Entomological Society of America*, 114(5):541–546, November 2020.

[Pasquier and Grüter, 2016] Grégoire Pasquier and Christoph Grüter. Individual learning performance and exploratory activity are linked to colony foraging success in a mass-recruiting ant. *Behavioral Ecology*, page arw079, June 2016.

[Sylvester and Gini, 2023] A. Sylvester and M. Gini. Facilitating real-time collaboration and learning in search environments for multi-robot systems via real-time evolutionary algorithm. In *ARMS Workshop at AAMAS*, London, England, 2023.