Resource-Aware Adaptation of Heterogeneous Strategies for Coalition Formation

Extended Abstract

Anusha Srikanthan
University of Pennsylvania
Philadelphia, Pennsylvania, USA
sanusha@seas.upenn.edu

Harish Ravichandar
Georgia Institute of Technology
Atlanta, Georgia, USA
harish.ravichandar@cc.gatech.edu

ABSTRACT

Existing approaches to coalition formation assume that task requirements are precisely specified by the human operator. Further, existing approaches ignore the fact that tasks could often be performed by following one of many equivalent strategies. However, prior work has demonstrated that humans, while extremely adept at solving complex problems, struggle to *explicitly* state the intuition that led to their solution. In this work, we propose a two-part framework to i) learn *implicit* heterogeneous strategies for coalition formation directly from expert demonstrations, and ii) *adaptively* select one of the inferred strategies based on available resources, without additional training.

KEYWORDS

Multiagent Systems; Coalition Formation; Heterogeneity

ACM Reference Format:
Anusha Srikanthan and Harish Ravichandar. 2022. Resource-Aware Adaptation of Heterogeneous Strategies for Coalition Formation: Extended Abstract. 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

Coalition formation problems require a team of agents to be partitioned into non-overlapping sub-teams (i.e., coalitions) in order to perform multiple concurrent tasks [1, 2]. In this work, we are interested in coalition formation for *heterogeneous teams* (i.e., teams made of agents with different capabilities). Formally, task allocation problems are categorized based on three axes: Single-Task (ST) vs. Multi-Task (MT) robots; Single-Robot (SR) vs. Multi-Robot (MR) tasks; and Instantaneous Allocation (IA) vs. Time-extended Allocation (TA). Our work addresses the coalition formation problem, which is an instance of the ST-MR-IA problem.

Most existing approaches to coalition formation assume that the requirements associated with different tasks are explicitly specified by the human operator (e.g., [5, 6]). However, prior work has demonstrated that manually specifying multi-dimensional objective functions that capture trade-offs between various factors can be very challenging [3, 8, 9]. Therefore, we extract implicit task requirements from expert demonstrations, similar in spirit to inverse reinforcement learning (IRL). We model task requirements in terms of *traits* (i.e., multi-dimensional capabilities) that are required to accomplish the task.

A rich body of work in psychology and human-robot teaming suggest that complex tasks are often solved using one of many comparable strategies [4, 7]. For instance, consider a search task as part of a disaster response mission. To search effectively, one can either allocate i) a slow-moving coalition with a large collective sensing area, or ii) a fast-moving coalition with a small collective sensing area. We denote such different, yet equivalent, trait requirements as *heterogeneous strategies.* Our approach explicitly accounts for and extracts such diverse strategies. Further, we propose a strategy-selection algorithm that chooses the most appropriate strategy given the resources available from a new team.

In summary, we contribute: 1) a coalition formation framework that accounts for heterogeneous strategies, 2) a clustering-based approach for inferring such generalizable heterogeneous strategies from expert demonstrations, and 3) an optimization-based method for resource-aware strategy selection that can generalize to entirely new teams without additional training.

2 PROBLEM FORMULATION

We consider a heterogeneous team composed on S *species* (i.e., robot types), in which the s^{th} species contains N_s robots. Let the team be tasked with a set of M concurrent tasks denoted by $\mathcal{T} = \{T_1, T_2, .., T_M\}$. Please refer to [10] for a more formal description.

Let a set of heterogeneous **strategies** associated with the m^{th} task be made up of P_m strategies. The r^{th} strategy for the m^{th} task is given by ${}^{r}y_m^* \in \mathbb{R}_+^U, \forall r \in \{1, \cdots, P_m\}$. Let a set of N expert **demonstrations** in the form of robot assignments be given by

$$\mathcal{D} = \{X^{(i)},\ Q^{(i)}\}_{i=1}^{N} \tag{1}$$

where $X^{(i)}$ represents the expert-specified assignment matrix for a team with capabilities encoded by the Species-Trait matrix $Q^{(i)}$.

Given the above definitions, our problem consists of two steps: *i)* extracting the set of approximated strategies ${}^{r}\hat{y}_m$ for each task from $\mathcal{D}$, and *ii)* optimizing the Assignment matrix $X^{(j)}$ of a new team with Species-Trait matrix $Q^{(j)} \notin \mathcal{D}$.

3 PROPOSED METHOD

3.1 Extracting heterogeneous task requirements

To extract implicit task requirements, we first compute the trait aggregations from the demonstrations $\mathcal{D}$ as $y_m^{(i)} = {Q^{(i)}}^T x_m^{(i)}, \quad \forall m, i$ where $y_m^{(i)}$ represents trait aggregation associated with the m^{th} task of the i^{th} demonstration. We apply agglomerative (i.e., hierarchical

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.

bottom-up) clustering on the computed trait aggregation vectors for each task to obtain clusters denoted by $\{{}^rC_m\}_{r=1}^{P_m}$. Once the clusters are identified, we compute the approximate task requirements associated with each strategy on every task as follows

$$ {}^r\hat{y}_m = \sum_{y_m^{(i)} \in {}^rC_m} \frac{y_m^{(i)}}{|{}^rC_m|} \tag{2} $$

where $|{}^rC_m|$ denotes the number of demonstrations for Task T_m that are identified as a part of the r^{th} cluster. Revisiting our disaster response example, we could extract two distinct strategies from multiple expert coalitions assigned to the search task: one requiring large sensing radius and slow speed, and another requiring small sensing radius and high speed.

3.2 Resource-aware coalition formation

To denote the choice of strategy for each task, we introduce one-hot encoded *strategy selectors* $z_m \in \{0,1\}^{P_m}, \forall m$. We also introduce integer decision variables $x_m \in \mathbb{Z}_+^S, \forall m$ representing the assignment of robots to each task. We simultaneously optimize the overall assignment such that the chosen set of trait requirements are satisfied for all tasks. Let ${}^re_m$ be the trait mismatch error between the aggregated traits and the trait requirements of the r^{th} strategy ${}^re_m = \|{}^r\hat{y}_m - Q^{(j)^T} x_m\|_2^2$ where x_m represents the assignment for the Task T_m. Hence, the net trait mismatch error for Task T_m is given by $E_m = z_m^T\, e_m$ where $e_m = [{}^1e_m, \cdots, {}^{P_m}e_m]^T \in \mathbb{R}_+^{P_m}$.

Finally, we cast the *resource-aware* optimization of robot assignments for the new team with $Q^{(j)} \notin \mathcal{D}$ as a constrained quadratic integer program:

$$ \{x_m^{*(j)}, z_m^{*(j)}\}_{m=1}^{M} = \arg\min_{x_m, z_m} \sum_m E_m \tag{3} $$

$$ \text{s.t.} \sum_m x_m \le N_a,\ z_m^T \cdot \mathbf{1} = 1,\ \hat{Y}_m z_m \le Q^{(j)^T} x_m,\ \forall m \tag{4} $$

where $N_a \in \mathbb{Z}_+^S$ represents the vector of total robots per species, $\mathbf{1}$ is a vector of ones, and $\hat{Y}_m = [{}^1\hat{y}_m, \cdots, {}^{P_m}\hat{y}_m] \in \mathbb{R}_+^{U \times P_m}$ represents all the distinct trait requirements for Task T_m extracted from the demonstrations.

Revisiting our disaster response example, our framework chooses between the two strategies for the search task (low speed and large sensing radius vs. high speed and small sensing radius) depending on the capabilities of the available team. Indeed, one of the two strategies is likely to be better suited than the other for a given target team. Further, the resource constraint in (4) helps our framework realize that if all ground vehicles are assigned to the search task, we will not be able to utilize them to remove debris.

4 EXPERIMENTAL EVALUATION

We evaluate our approach against several baselines, including some that resemble existing approaches, using detailed numerical simulations, StarCraft II battles, and a multi-robot emergency-response scenario. Our results indicate that our framework consistently outperforms all baselines in terms of requirement satisfaction, resource utilization, and task success rates. For further details and more comprehensive results, please refer to [10].

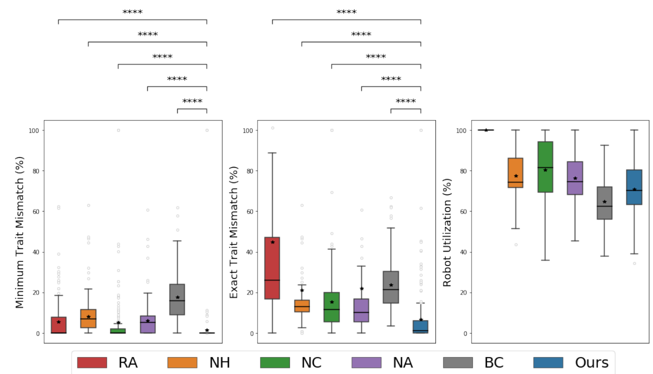

Figure 1: The figure shows subplots of measures of minimum trait mismatch, exact trait mismatch, and robot utilization (from left to right) observed in the numerical analyses. As shown, our approach achieves the lowest minimum and exact trait mismatch percentage error computed across 3 tasks over 60 test teams. We performed the Kruskal-Wallis test, followed by the Dunn test for post-hoc pairwise comparisons and FDR adjustment, and found that all comparisons are statistically significant ($p < 1e-5$).

4.1 Numerical analyses

We evaluate our approach against four baselines using numerical simulations across a wide variety of problems by altering aspects such as team size, robot capabilities, and task strategies. We consider simulated task allocation problems, each with four species ($S = 4$), three traits ($U = 3$), three tasks ($M = 3$), three strategies per task ($P_m = 3, \forall m$), and the number of robots per species uniformly randomly sampled between 6 and 33. We generate teams and strategies such that our data contains a mix of under-, sufficiently-, and over-resourced teams. We set 1800 seconds as an upper bound on the optimization time for all approaches.

Our approach identified three distinct strategies for each task from the demonstration set $\mathcal{D}$. Given the inferred strategies, our approach outperformed all baselines pointing to the deficiencies of existing coalition formation approaches that either ignore heterogeneous strategies or the context of available resources (see Fig. 1). Moreover, the results indicate that the baseline embodying unstructured supervised learning performed the worst, as it does not account for inter-task dependencies nor does it generalize to any new species.

4.2 Evaluation on StarCraft II

We designed battles on the game, StarCraft II, to emulate tasks that require careful allocation using combinations of the available species. The key finding from this experiment was the fact that ignoring heterogeneity and relying on statistical averages of multimodal distributions could lead to adverse effects.

ACKNOWLEDGMENTS

This work was supported by the Army Research Lab under Grants W911NF-20-2-0036 and W911NF-17-2-0181 (DCIST CRA).

REFERENCES

[1] Brian P Gerkey and Maja J Matarić. 2004. A formal analysis and taxonomy of task allocation in multi-robot systems. *The International journal of robotics research* 23, 9 (2004), 939–954.

[2] G Ayorkor Korsah, Anthony Stentz, and M Bernardine Dias. 2013. A comprehensive taxonomy for multi-robot task allocation. *The International Journal of Robotics Research* 32, 12 (2013), 1495–1512.

[3] Richard E Nisbett and Timothy D Wilson. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological review* 84, 3 (1977), 231.

[4] John W Payne, James R Bettman, Eloise Coupey, and Eric J Johnson. 1992. A constructive process view of decision making: Multiple strategies in judgment and choice. *Acta Psychologica* 80, 1-3 (1992), 107–141.

[5] Amanda Prorok, M Ani Hsieh, and Vijay Kumar. 2017. The impact of diversity on optimal control policies for heterogeneous robot swarms. *IEEE Transactions on Robotics* 33, 2 (2017), 346–358.

[6] Harish Ravichandar, Kenneth Shaw, and Sonia Chernova. 2020. STRATA: A unified framework for task assignments in large teams of heterogeneous agents. *Autonomous Agents and Multi-Agent Systems* 34, 2 (2020).

[7] Paul B Reverdy, Vaibhav Srivastava, and Naomi Ehrich Leonard. 2014. Modeling human decision making in generalized Gaussian multiarmed bandits. *Proc. IEEE* 102, 4 (2014), 544–571.

[8] Jörg Rieskamp, Jerome R Busemeyer, and Tei Laine. 2003. How do people learn to allocate resources? Comparing two learning theories. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29, 6 (2003), 1066.

[9] Jörg Rieskamp and Philipp E Otto. 2006. SSL: a theory of how people learn to select strategies. *Journal of Experimental Psychology: General* 135, 2 (2006), 207.

[10] Anusha Srikanthan and Harish Ravichandar. 2022. Resource-Aware Adaptation of Heterogeneous Strategies for Coalition Formation. arXiv:2108.02733 [cs.MA]