Check for updates

Proceedings of the ASME 2024
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC-CIE2024
August 25-28, 2024, Washington, DC

DETC2024-143326

IMPACT OF SOCIAL LEARNING ON TEAMWORK EFFICIENCY IN LEARNING-BASED SELF-ORGANIZED SYSTEMS

Bingling Huang*
Dept. of Mechanical Engineering
California State University, Fullerton.
Fullerton, USA
bihuang@fullerton.edu
(*Corresponding Author)

Hao Ji
Center for Advanced Research Computing
University of Southern California
Los Angeles, USA
haoji@usc.edu

Yan Jin
Dept. of Aerospace & Mechanical Engineering
University of Southern California
Los Angeles, USA
yjin@usc.edu

ABSTRACT

Self-organized systems (SOSs) possess the great potential to adapt to dynamic environments, and thus, advance the level of autonomy in engineering systems. Learning-based approaches, particularly multiagent reinforcement learning (MARL), have empowered the design freedom of SOSs while enhancing the system to be more intelligent. This is primarily because MARL encourages agents to learn from past failures and data through training. However, the training process is conducted in a "black box", whose embeddings are uninterpretable for humans yet. This poses significant challenges for system designers, especially when task complexity reaches a certain extent, and agents experience a shortfall in learning due to limited available information. Therefore, in order to address how to steer more effective learning, we broaden the views of individual agents, enabling them to recognize the presence and actions of other members within a certain range of distance. This approach fosters social learning within AI agent groups. Prior work, however, has shown the intricacy of the social mechanism. Inappropriate social abilities can undermine a team's performance. Thus, it is imperative to delve into the effects of social learning to facilitate the system design, further narrowing down the understanding gap between humans and AI agents. In this study, we train agent teams to engage in social learning to complete an assembly task involving collision avoidance. We aim to answer how social abilities influence teamwork at both the team and individual agent levels in terms of energy utilization. By examining the impact of team energy efficiency and teamwork division, our findings indicate a mapping linking appropriate social abilities to tasks of a certain complexity level. Suitable social abilities contribute to elevated team intelligence, more effective teamwork division, and higher energy efficiency, while insufficient social abilities may hinder cooperation and reduce energy efficiency.

Keywords: Self-organized systems, social learning, design, multiagent reinforcement learning, teamwork efficiency.

1. INTRODUCTION

Robotic collaboration is increasingly recognized as the key to achieving high autonomy in the realm of engineering [1]. For example, robot teams can be deployed for repetitive tasks in production lines [2], or hazardous environments in rescue missions [3], taking over the traditional work performed by humans. Achieving such autonomy demands more from system design, especially as tasks grow in complexity [4]. To address this issue, self-organizing approaches have been applied widely thanks to their high-level adaptability [5]. Those systems possess abilities to change their internal structure and functions in response to changing external circumstances [4], therefore greatly alleviating the challenges associated with human design.

Learning-based approaches, particularly multiagent reinforcement learning (MARL) [6], have been applied to aid the design of SOSs, attaining a high degree of robustness, intelligence, and capabilities. However, while this approach offers various advantages such as high flexibility and minimum needs for the top-down system design [7], it also introduces issues in implementation due to its abstract nature. What makes the MARL approach more challenging is how to guide effective team learning

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

with limited available information in a dynamic and unknown environment.

Our research investigates strategies to steer effective training under these demanding conditions. Humans can gain incredible capabilities through social interactions with other people [8], a phenomenon that happens across human society and animal groups, which benefits groups' adoption and evolution [9]. This insight brings a natural question in SOS design, "*could social learning facilitate the intelligence growth of AI robot groups?*"

Driven by motivation, our previous work has integrated social learning in a MARL-based system [10], revealing its big potential to enhance learning effectiveness. However, the intricacy of social mechanisms within AI agents is evident, where inappropriate social abilities can undermine team performances [10]. What's more, in the context of engineering applications, an ideal agent team should not only output desirable performances but also employ efficient teamwork strategies that make reasonable use of resources. Those put higher requirements on the system design. Therefore, it is imperative to delve deeper into the social learning mechanisms to guide the system design and narrow the gap of understanding between human designers and AI agent teams.

To address these issues, this study concentrated on exploring the mechanism of social learning at both team and individual agent levels. We seek to investigate how social learning influences teamwork efficiency by examining team energy utilization and teamwork division among the AI team. In the context of high-precious assembly tasks involving collision avoidance, we probe two research questions, 1) *how do social abilities influence energy efficiency at the team level?* and 2) *how do social abilities impact the teamwork division at the individual agent level?* Our goal is to offer insights that enable designers to effectively develop a learning-based multiagent system, ensuring predictable and desirable outcomes, and reasonable use of resources.

The rest of the paper is organized as follows: Section 2 presents related work on this topic. Section 3 illustrates the employed methodologies. Section 4 describes the experimental design and settings. Section 5 offers the experimental results along with a thorough analysis and discussions. Finally, the conclusion and future direction are presented in Section 6.

2. RELATED WORK

Social learning theory [8] proposed by Albert Bandura has emphasized the importance of observing, modeling, and imitating the behaviors, attitudes, and emotional reactions of others. The learning abilities obtained from interaction inspire researchers to explore the ways in which social learning can augment the intelligence of the agent groups.

In [11], the authors have delved into the social learning mechanisms within the context of collective robot exploration. The authors examined how social learning influences the of emergent group behaviors, conducting simulations where the agents follow predefined exploration trajectories including spiral, boustrophedon, and sawtooth patterns [11]. The performances of asocial teams (without interaction), social teams (with interaction), and combined teams are compared. It is found that deterministic asocial exploration is best performed by a spiral exploration mechanism, but its effectiveness diminishes as robots share social rewards [11]. Besides, continuous sharing of social rewards provides the most optimal outcomes, though it demands a significant energy investment [11]. However, given that each agent acts independently and prioritizes its own best interests, minimal cooperation is required for this foraging task. Additionally, the use of predetermined exploration strategies constrains the potential applications in engineering fields.

In [12], the authors have proposed an integrated approach by combining individual learning and social learning to solve the coordination problem in multiagent systems. This allows agents to switch between these two learning modes during task execution. The results identified the distinct roles for different learning strategies, suggesting that social learning is less valuable than individual learning in open and dynamic environments due to the uncertainty introduced by social information [12]. This highlights a challenge in leveraging social learning effectively. Given that the MARL is inherently a dynamic learning process, people should be cautious about incorporating additional information. Previous research supports this, indicating that inappropriate social information can negatively impact team performance [10]. Nonetheless, our goal is to intelligently design systems and effectively utilize resources to overcome these issues.

In [13], the authors have investigated the multiagent coordination problems from the systematical perspective. The authors examined two types of cooperative learning, individual action learners, and joint action learners, across scenarios with different factors, including topology networks, topology parameters and types of games [13]. It reveals that joint action learners that conduct both individual learning and learning from neighbors outperform individual action learners [13]. The results provide valuable insights that the network topology types matter in team performances, even in the low-dimension state space games.

In [14], the authors address a social learning problem within a network where agents make decisions in sequence. It delved into the difference between two types of agents, rational and irrational agents, depending on whether they take into account the decisions made by others [14]. The studies highlight the application of Bayesian theory by the agents to navigate the decision-making process.

There is research focused on incorporating social elements into reward functions to facilitate learning. Borsa et al. [15] demonstrate the potential of pure observations. It is shown that novice agents can conduct learning through observing the behaviors of "teachers". However, the reliance on a "teacher" agent poses challenges to providing a model for imitation, which is demanding for complex engineering tasks. Ideally, an agent team can develop a good teamwork strategy through self-learning and self-evolving [10]. On one hand, identifying suitable predefined strategies is challenging. On the other hand, such strategies may prevent agents from discovering optimal solutions and team policies.

In [16], the authors proposed an auxiliary loss to enhance the team behaviors. The loss encourages novice agents to learn from "experts" by minimizing the discrepancy. In this way, novice agents can refine their strategies towards "expert" state trajectories. Rather than compelling agents to directly copy expert

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

actions, the method employs state demonstrations as a guide, offering a more practical approach.

While previous studies have provided aspects of social learning, few explicitly address its application in engineering. Given that multiagent systems are sensitive to input information, it is imperative to enhance our comprehension of social learning mechanisms from a systematic perspective. Such knowledge would enable the applications to solve complicated engineering tasks effectively and intelligently.

3. METHODS

3.1 Multi-agent Reinforcement Learning

Multi-agent reinforcement learning (MARL) has been broadly utilized to model and solve the tasks completed by more than one agent [17]. We focus on the cooperative MARL to explore multiple agents acting, interacting, and influencing the environment simultaneously in the context of complex engineering tasks. The cooperative MARL is modeled in a tuple of $< S, A, P, R, \gamma, N, t >$ in a sequence of discrete-time steps [18]. In the modeling tuple, t represents the time step; N signifies the number of agents; S stands for the information of the environment; A is the representation of the joint actions of all agents and a_t^i stands for the ith agent's action choice at time t. Besides, the transition function is given by $P(s_{t+1}|s_t, a_t): S \times A \times S \rightarrow [0, 1]$ for modeling the environment and determining the environmental states. The system actions are defined as a joint set $A = A^1 \times \cdots \times A^N$ [6]. The joint policy presents in ${}^{S}\Pi \sim {}^{S}\Pi^1 \times {}^{S}\Pi^2 \times \ldots \times {}^{S}\Pi^N$, pointing at all agents' learned policies. All agents share a team reward function represented by R for a common goal in the context of collaborative tasks. The reward functions use numerical scores to assess the team performance. Besides, the discount factor, γ, is employed in computing return values, reflecting the diminishing influences of the future effect on the present value. Overall, the problem of solving an MDP is to find an optimal policy π that can maximize the accumulated reward [19].

Deep Q-learning algorithm [20] is applied in our research, taking advantage of its powerful abilities of deep learning. In the context of individual learning, each agent perceives its peers as part of the surrounding environment, leveraging the Deep Q-learning framework and their own neural networks in an end-to-end way [21] [22]. The loss function is shown in equation (1), working for iterative update of Q-network's weight θ_j^i of each agent i at each iteration j.

$$L_j(\theta_j^i) = E\left[\left(r_{t+1} + \gamma \max_{a^i} Q\left(s_{t+1}, a_{t+1}^i \mid {}^{-}\theta_j^i\right) - Q\left(s_t, a_t^i \mid \theta_j^i\right)\right)^2\right] \quad (1)$$

where θ_j^i is the parameters of the target networks and ${}^{-}\theta_j^i$ is the parameters of the online network.

In the cooperative task domains, agents are trained aiming to maximize cumulative rewards derived from their collective actions over time. Their individual neural networks are trained independently and concurrently to capture the cooperative behavior codes in a dynamic environment.

3.2 Integrate Social Ability in Multi-agent Teams

Many studies have shown the limitation of individual learning [23] [24], where ignoring other members' existence tends to hinder teamwork. To enhance and boost the capability of an agent team, it is crucial to examine the role of social abilities within multi-agent systems.

In a team, agents conduct learning based on the information available to them. In independent learning, agents only receive task-relevant information, such as the environmental features, task progress and object information, as shown in Fig. 1. However, the information is usually insufficient to conduct complex tasks or to coordinate a large team [10]. To bridge this gap, the concept of *social views* is introduced, encompassing interactive information like the agent's ID, name, plans, etc. With access to social views, agents are able to conduct social learning [10] and have the potential to output better teamwork [22].

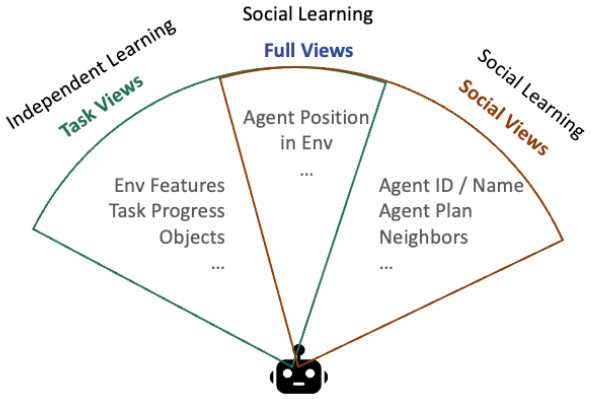


Figure 1: *Task views and social views of a robot.*

To incorporate social information in multiagent reinforcement learning, we define a distinct component of the state space, named as social state ${}^{K}S$. This state reflects the social capabilities of the robots. On the other hand, the task-relevant information is saved in the environmental state ${}^{E}S$. They are appended into a unified state ${}^{S}S$, shown in equation (2), that is utilized throughout the training and the iteration.

$${}^{S}S = {}^{E}S \frown {}^{K}S \quad (2)$$

Therefore, the loss function for ith agent is updated as below:

$$L_t^i({}^{S}\theta_t^i) = E_{s_{s,a,r,s'}}\left[\left(r_t^i + \gamma \max_{a'} Q\left({}^{S}s_t^{i\prime}, a' \mid \theta_{t-1}^i\right) - Q\left({}^{S}s_t^i, a \mid \theta_t^i\right)\right)^2\right] \quad (3)$$

In this study, the social ability is referred to "*the capacity to recognize the presence and conditions of fellow team members*". This concept aims to mimic real-world scenarios where robots utilize sensors to send and receive signals, enabling them to communicate and share information effectively [25].

3.3 Multi-agent Teams with Different Social Abilities

Since the social abilities in this paper are limited to observing other members' action choices without differentiating their

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

identities, Thus, the social state ${}^{K}S$ can be represented in a tuple holding the observed signals. The dimensions of ${}^{K}S$ varies depending on different observational ranges, which simulates the scenarios where robots can perceive signals from their neighboring robots.

To comprehensively assess how social abilities influence team dynamics, this research delineates distinct levels of social capabilities within the framework of assembly tasks:

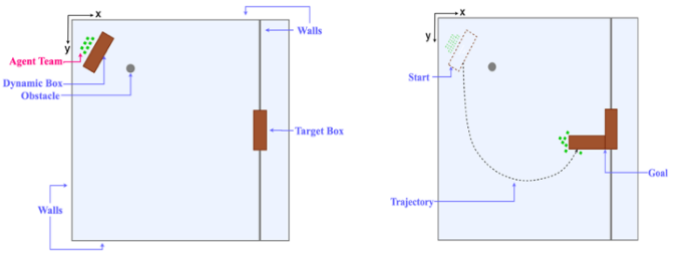
- *No Social (NS) – Agents lack any form of social abilities.*
- *Observe Neighbors (ON) with different vision ranges – Agents have the capability to monitor the adjacent areas of their current position, determining whether these areas are influenced by the actions of other agents.*

4. EXPERIMENT DESIGN: 'L'-SHAPE ASSEMBLY INVOLVING COLLISION AVOIDANCE

4.1 Task Description

Precision assembly occupies an important position in the engineering field [26]. For instance, products are constructed assembled from parts in manufacturing production lines. During the process, it is desired not only to assemble products with high precision as well as avoid any collisions to ensure safety. Consequently, our study will concentrate on an "L"-shape assembly involving collision avoidance tasks to simulate the scenarios encountered in factory settings.

Figure 2 depicts the "L-shape" assembly tasks involving collision avoidance, where an agent team is tasked to transport a Dynamic Box and assemble it with a Target Box to form a "L" configuration. It illustrates the task start and goal status as shown in Fig.2-(a) and Fig. 2-(b), respectively. During the task process, it is essential for the Dynamic Box to be controlled without contacting with any surrounding walls or obstacles in the field.



(a) Task start status (b) Task goal status

Figure 2: *Task environment introduction*

The task environment is designed as a 1000×1000 pixels field, comprising a movable Dynamic Box, a stationary Target Box, stationary obstacles and four surrounding walls. The Dynamic Box (depicted as the upper-left brown rectangle) can be moved by agent teams and has a mass of 1 kg. The Target Box is positioned at the center of the right boundary and remains static throughout the task. Both the Dynamic and Target Boxes are 180×60 pixels. The agents, represented as green squares, need to spontaneously organize themselves to transport the Dynamic Box through an appropriate trajectory to assemble two boxes into an "L" configuration, shown as Fig.2-(b). At the same time, they are required to ensure that the Dynamic Box does not collide with the walls or any obstacles.

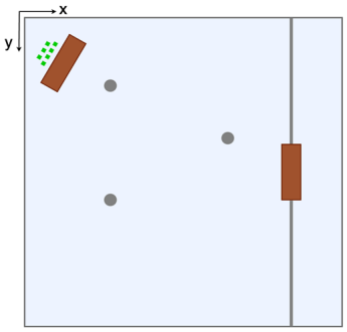


Figure 3: *Task description*

Our study focuses on a complex assembly task characterized by three obstacles, as shown in Figure. 3, with specifics provided in Table 1. This task's complexity is heightened due to the limited pathways for the Dynamic box hindered by three obstacles. Agent teams need to know how to transport and rotate the Dynamic box collaboratively to avoid collisions while successfully assembling an "L"-shape. Meanwhile, the box's movement physically depends on the collective impulses provided by all agents at each step. Those conditions impose rigorous demands on team intelligence and the strategic coordination of group policies. Hence, achieving high group intelligence is critical for the success of the task, which requires the effective use of available information and proper system design.

Given the inherent properties of deep learning, the training process is highly sensitive to its input data [33]. Consequently, the quantity and representation of social information are crucial to the success of the training. This paper will thus explore the impact of social abilities on team performances and delve into the learning mechanism of a multiagent team. It should be noted that the tasks are simulated by *pygame* [27] environment.

Table 1: *Task settings*

Field size	1000 * 1000
Box size	180 * 60
Target box size	180 * 60
Box start center coordination	(150, 180)
Target box center coordination	(950, 500)
Obstacle center coordination	(350, 170) (350, 370) (500, 270)
Box mass (kg)	1
Push impulse (1 $N \cdot s$)	1

4.2 Algorithm, Action Space and State Space

The agent teams are trained using Deep Q-learning [20], incorporating an epsilon-greedy (ε–$greedy$) strategy [28] to facilitate the exploration of the optimal policy. Table 2 lists the hyperparameters employed in the algorithm, including training episodes, discount factors, learning rates, and others.

Table 2: *Hyperparameters of training*

Training episodes	20,000
Discount factor	0.99

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

Memory buffer size	1000
Mini-batch size	32
Target network update frequency	200
Learning rate	0.001
Neural network size	(63, 64, 128, 6)
epsilon	1 → 0.01

Action Space: Within the context of assembly tasks, the agent teams are capable to transport the Dynamic Box through pushing. Therefore, we divide the surrounding area of the dynamic box in six regions, shown in the Figure 4. During the simulation process, each individual agent needs to make a choice to exert a pushing force equivalent to 1 N·s impulse on the box. Given the relatively small size of each agent, it is feasible for several agents to occupy a single region concurrently at any steps. This arrangement allows for the coordination and application of collective forces.

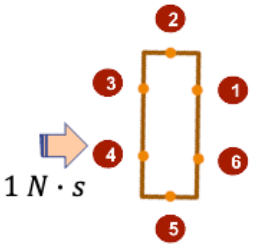


Figure 4: *Action space*

Therefore, the action space is established as follows:

$$A = < a_1, a_2, a_3, a_4, a_5, a_6 > \quad (4)$$

where a_i *means push i position with impulse of* $1\ N \cdot s$

State Space: The state space consists of two parts, the *environmental state* which provides information on the surrounding environment, and the *social state* which includes the fellow agents' statuses in a certain range.

The *environmental state*, represented as ${}^{E}S$ is set as a 63-digit tuple, shown as in equation (5). This tuple encompasses data regarding the vicinity situation around the Dynamic Box within its first 60 digits, while the last three digits details the Dynamic Box's velocity in x and y directions and its angular velocity. Thus, the *environmental state* can be defined as:

$${}^{E}S = < Vicinity\ Situation, v_X, v_Y, \omega > \quad (5)$$

where *Vicinity Situation* is encoded as a tuple of binary number ("0" for absence or "1" for the presence), indicating if there exists obstacles / target box around the Dynamic Box. This is determined through sensors equipped on the Dynamic Box, which have a detection range of 200 pixels.

The *social state*, denoted as ${}^{S}S$, is represented in a tuple to reflect the social dynamics among team members. Given the action space setting, where an agent occupies one of six designated regions to apply force, the social state allows agents to discern the presence of other team members in certain regions. We signify this presence as an indicator I, where $I == 1$ means that at least one team member is exerting force in that region, and $I == 0$ indicates the absence of such force. Therefore, the *social state* varies based on the level of social abilities in the agent team.

Building upon the social abilities outlined in section 3.3, we further detail the social abilities as follows in the context of the assembly tasks.

- *Observe Two Neighbors (ON2) – Agents have the capability to monitor the immediate adjacent 2 areas to them, determining whether these areas are influenced by the actions of other agents.*
- *Observe Four Neighbors (ON4) – Agents have the capability to monitor the immediate adjacent 4 areas to them.*
- *Observe Five Neighbors (ON5) – Agents have the capability to monitor the immediate adjacent 5 areas to them.*

In the case of "Observe Two Neighbours (ON2)", we define that one agent (shown as symbol *robot* in Fig. 5-(b)) has the capacity to recognize the presence of other agents in its immediately adjacent left and right adjacent areas (shown as the symbol "√" in Fig. 5-(b)). Thus, the *social state* is defined as:

$${}^{S}s_i = < I_{left}, I_{right} >_{ON2} \quad (6)$$

where I_{left} and I_{right} serves as indicators to show whether the presence of team members applying force in the left and right adjacent areas of the ith agent.

Similarly, when the agent has broader views to observe its immediate left two and right areas shown in Fig. 5-(c), the *social state* in the case of "Observe Four Neighbours (ON4)" is represented by a four-digit tuple. In this case:

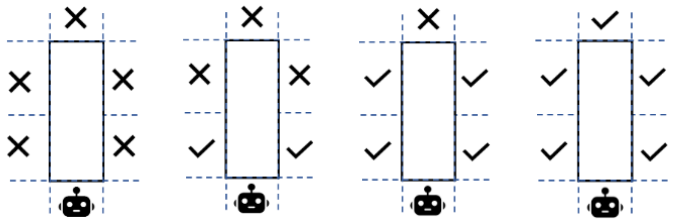
$${}^{S}s_i = < I_{left2}, I_{left1}, I_{right1}, I_{right2} >_{ON4} \quad (7)$$

where I_{left1} and I_{right1} indicate force application in the direct left and right adjacent areas of the ith agent, while I_{left2} and I_{right2} point to the secondary adjacent areas.

In the case of "Observe Five Neighbours (ON5)" shown in Fig. 5-(d), one agent can see the presence of others in all surrounding positions wherever it stands, shown in Fig. 5-(d). The *social state* is thus represented as:

$${}^{S}s_i = < I_{left2}, I_{left1}, I_{opposite}, I_{right1}, I_{right2} >_{ON5} \quad (8)$$

where I_{left1}, I_{right1}, I_{left2}, and I_{right2} indicate the presence of team members in both the direct and secondary adjacent areas, with $I_{opposite}$ pointing to the opposite area.



(a) NS (b) ON2 (c) ON4 (d) ON5

Figure 5: *Agent with different social abilities.*

Consequently, the overall state is constructed as ${}^{S}S = {}^{E}S \frown {}^{K}S$, integrating data by concatenating the *environmental state* and *social state*. This leads agents to learn and iterate their policies from both task-related and social-related signals. It is noted that all agents within a team are homogeneous, therefore they receive the identical environmental states and possess the same social abilities. However, the specific social information available to

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

each agent may differ, influenced by their relevant positions to the Dynamic Box and the viewings.

When agents receive solely the *environmental state*, the team is considered to lack social capabilities. In our study, this setting serves as the baseline and is named as No Social (NS), as depicted in Fig. 5-(a).

4.3 Measurement: Team Performance Score, Team Energy Efficiency, and Teamwork Division

4.3.1 Team Performance Scores (PS)

Task performance is a key indicator of the group intelligence, assessed through task quality and completeness. In our research, we define and utilize Performance Scores (PS) as a means to evaluate task success, focusing on the final configurations of the two boxes.

The critical aspect of our evaluation is the precision of the "L"-shape achieved by the agent team. We track two specific angles to identify the two boxes' positions: the self-angle of Dynamic Box: α and the relevant angle between the Dynamic and Target Boxes: β. They both can be detected by sensors on the Dynamic Box, as shown in Figure. 6. The reasons for selecting those two angles are, firstly, the two angles can fully define the relative positions of the two boxes, and secondly, they are sensible and easily interpretable signals.

Thus, the final configuration can be determined by measuring the signals α_T and β_T at the termination episode, shown in Fig. 6-(b) and formulated in equation (7).

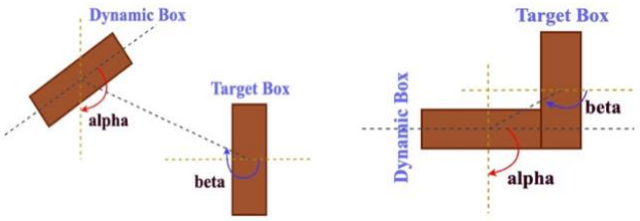


Figure 6: *Final configuration in terms of alpha and beta angles.*

$$Final\ configurations = < \alpha_T, \beta_T > \quad (9)$$

where T *is the termination episode.*

Task performance is then evaluated through comparing the trained final configurations and the ideal configuration. Geometrically, an ideal "L-shape" is characterized by α = 90° and β = 135°, as represented in Figure 6-(b). Based on this, we set two hyperparameters for the goal of signal self-angle α and the relevant angle β, as

$$\alpha_{goal} = 90° \quad (10)$$
$$\beta_{goal} = 135° \quad (11)$$

Therefore, we can define and calculate PS as shown in equation (12), offering a quantitative evaluation of the geometric configuration that aligns with the reward function [21].

$$PS = 100 \cdot exp(-|\alpha_T - \alpha_{goal}|) \cdot exp(-|\beta_T - \beta_{goal}|) \quad (12)$$

What's more, it indicates the degree to which the agent team achieves the systematical preferences. By setting the coefficient as 100, we ensure that PS falls within the range of 0 to 100. When a team is graded as 100, it means that the team has carried out a perfect outcome. This approach provides a clear and direct assessment of a team's learning ability and performance, as well as allowing categorizing the training results by setting a threshold for further analysis.

4.3.2 Team Energy Efficiency

In the context of robotics and autonomous systems, energy efficiency refers to the ability to optimize and minimize energy consumption while achieving desired objectives [29]. For a robot team, it involves maximizing the utilization of available energy resources, minimizing energy waste, and employing efficient strategies for performing tasks [30]. In practice, an energy-efficient robot team can reduce power consumption, prolong battery life, and maintain sustainable operations [30].

Thus, evaluating the energy efficiency of a team becomes essential, especially in a resource-constrained environment. That drives us to focus on the team energy efficiency of a well-trained agent team in performing complex assembly tasks, which aims at assessing whether the learned team policies can optimize both the team performance and energy consumption.

In a multi-robot team, each robot consumes a certain amount of energy to support its policy, such as optimal trajectory, stop times, motion strategies, and communication costs [31]. In our study, our focus is primarily on the cost associated with motion strategies, as the majority of energy consumption is attributed to the "push" motions executed by each robot. Thus, the energy consumption of a multi-robot team can be estimated by measuring the work performed by all robots, and each robot's energy consumption can be estimated roughly proportional to the Number of Steps (NOS) taken. As the team size remains constant throughout our subsequent studies, we employ Number of Steps (NOS) as a metric to evaluate energy consumption.

Consequently, the Performance Scores (PS) alongside the Number of Steps (NOS) can be integrated to evaluate team energy efficiency. This composite metric can be incorporated into a team energy efficiency evaluation matrix, as shown in Table 3.

Table 3: *Team energy efficiency evaluation matrix*

		Performance Scores (PS)	
		Low	High
Number of Steps (NOS)	Large	Poor	Mediocre
	Small	Mediocre	High

An ideal team is expected to complete a task in fewer steps while achieving good performance, meaning attaining optimal performance and economic efficiency simultaneously. Hence, the best case is when a team can achieve high PS within fewer NOS, implying that a team employs an efficiency strategy and minimizes energy consumption. Thus, it is considered "High" team energy efficiency. When a team obtains low PS with fewer steps, it means that the team has not wasted much energy on the task when dealing with a challenging task. In this case, we consider the team's energy efficiency as "Mediocre" even though the team is underperformed. One situation that can arise is when a team completes a task with

Copyright © 2024 by ASME

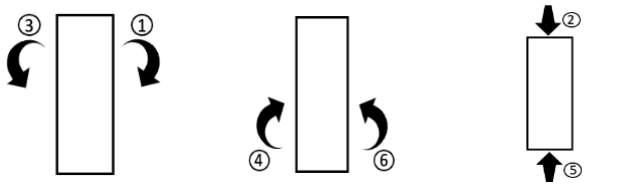
Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

a moderate performance score but requires a higher number of steps. This indicates that the team has to expend more energy to achieve the desired task output. Such team strategies can be considered as having "Mediocre" team energy efficiency, which is acceptable but not the most optimal. The worst cases are when teams fail to achieve both high task performance and reasonable energy utilization. That is classified as “Poor” team energy efficiency.

4.3.3 Teamwork Division

In addition to examining the team energy efficiency, we also analyze the workload division within a team at the agent-level. In a collaborative task, each robot incurs a certain cost to perform its role. It is crucial that each unit of cost results in a positive outcome to ensure efficient teamwork.

As known previously, six different actions can be selected by each robot to push one position around the Dynamic Box as shown in Fig. 4, and their joint actions result in the dynamic box’s movement, including displacement and rotations. Physically, there exist three pairs of actions that counteract each other, resulting in zero composition of forces or moments when taking the dynamic box center as the center of mass of the rigid body, as shown in Fig. 7.



(a) Action 1-3 (b) Action 4-6 (c) Action 2-5

Figure 7: *Conflicted action pairs.*

Fig. 7-(a) and (b) show that action pairs, Action-1 – Action-3 and Aciton-4 – Action-6, resulting in zero composition of external moments, while the pair of Action-2 – Aciton-5 leads to zero composition of external forces in Fig. 7-(c). Thus, if any two agents select conflicting action pairs, it is reasonable to say that they do not make any positive contributions to pushing the dynamic box at the current step. This situation should be avoided. For instance, if all agents engage in conflicting action pairs, this can result in energy consumption without meaningful output. In our studies, we assume that the team size is set appropriately so that all agents are expected to actively participate and make positive contributions to the task.

Hence, we consider any agents involved in the conflicted action pairs shown in Fig. 7 as *ineffective agents*, i.e., robots that are not currently contributing to the task. Meanwhile, the agents who contribute to the object’s movement at the current step are categorized as *contributors*, as shown in Table 4.

Table 4: *Teamwork division.*

	Explanation
Contributors	Those result in the motions of Dynamic Box.
Ineffective agents	Those involve in conflicted action pairs.

Therefore, the teamwork division can be assessed by examining agents’ action choices made at every step, allowing varying work allocations across different timesteps.

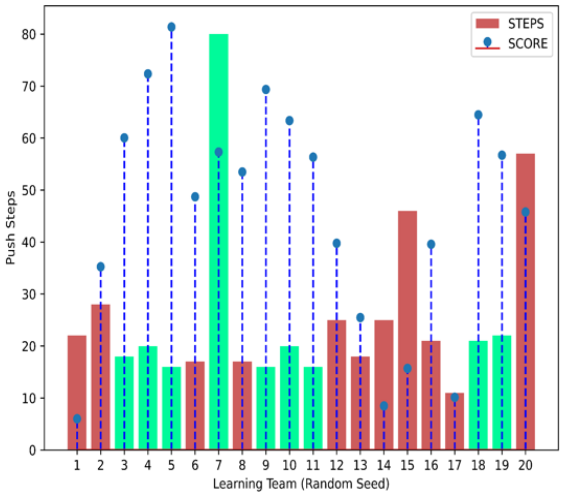
5. RESULT AND DISCUSSION

5.1 Teamwork Efficiency

Within a robot team, teamwork efficiency centers on the ability of these robots to work together effectively and harmoniously toward a common goal [7]. Drawing on principles from SOSs, agents within such a system maintain a degree of independence and autonomy [32]. Hence, their behaviors are often evaluated based on the teams’ overall performance, meanwhile, it is anticipated that the resources required for the task are in a feasible range. Thus, teams with higher levels of collective intelligence are recognized for their capability to deliver strong performance with strategic energy consumption, as explained in section 4.3.2. The challenge resides in achieving such a level of intelligence via effective training and learning. Based on the previous studies that highlighted the necessity of social learning in accomplishing complex tasks [10] [22], this study will delve deeper into the impact of social ability on teamwork efficiency to advance the understanding of social mechanisms.

In the current task, we assume that each action, defined as “*pushing the box with an impulse of 1* $N \cdot s$”, expends one unit of energy. The goal for teams is to achieve higher PS within a reasonable NOS, indicating a high team energy efficiency, as shown in Table 3.

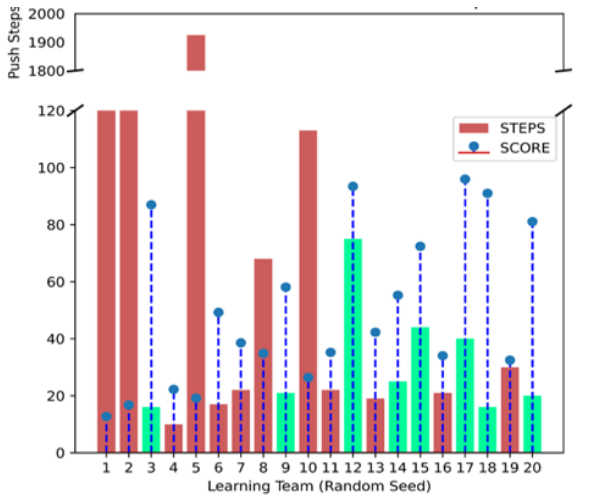
Therefore, to investigate the impact of social ability on sustainable operations and the expected task performance, we present a graphical analysis of PS and corresponding NOS for teams possessing different social abilities in Fig. 8. Subplots in Fig. 8 feature NS, ON2, ON4 and ON5 teams’ performances respectively, with Fig 8-(a) serving as a benchmark for comparison. Here, the x-axis is the team number (controlled by 1-20 random seeds). PS are depicted as blue dashed dots, while NOS are represented via bars. For the bar plots, red bars denote failure cases (where scores are less than 55), and green bars represent successful cases (where scores are larger than 55.)



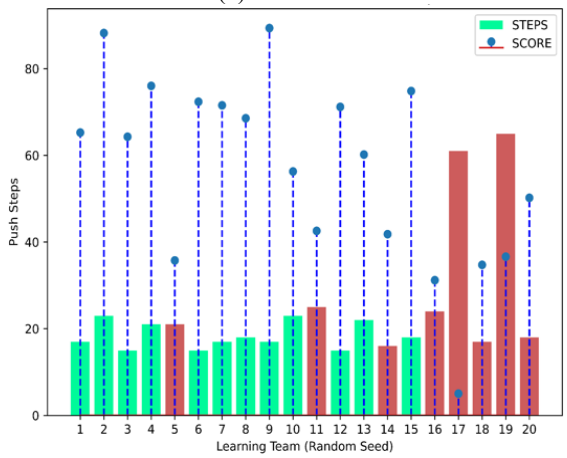
(a) NS Teams

Copyright © 2024 by ASME

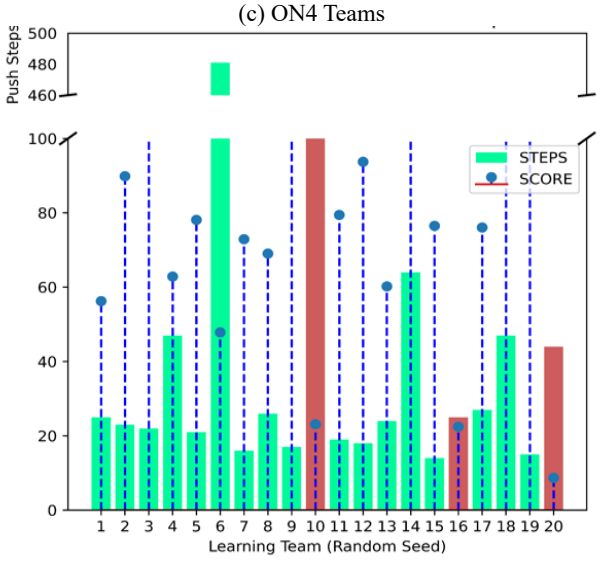
Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026



(b) ON2 Teams



(c) ON4 Teams



(d) ON5 Teams

Figure 8: *Team performance scores and number of steps*

Comparing the four subplots in Fig. 8, it can be seen that the impact of social ability on the team energy efficiency varies.

No Social Ability. In Fig. 8-(a), it can be observed that the teams without social abilities can hardly complete the complex task, achieving a mere 45% successful rate (9 successful cases out of 20). Nonetheless, the NOS required remains in a reasonable range, with most attempts concluding within 30 steps. This suggests that such teams are constrained in their capacity for exploration during the training, limited to environmental information alone. Consequently, while their ability to tackle complex tasks is restricted, they do not expend excessive energy on unsuccessful attempts.

Insufficient Social Ability. By incorporating teams with limited social abilities, such as ON2 with a restricted observational range, training results exhibit a higher failure rate compared to the benchmark. It can be observed that a significant portion of the ON2 teams spends considerable energy yet fails to meet the performance expectation. On one hand, ON2 teams show increased energy usage in the pursuit of optimal team policies, indicated by the greater number of required steps taken. On the other hand, energy usage does not correspond to higher returns in the form of task performance. As shown in Fig. 8-(b), not only PS are lower, but the success rate also drops to 40%.

The results suggest that the incorporation of social abilities has cost, potentially leading to increased energy consumption without guaranteed performance gains. For instance, the fifth ON2 team, represented by the fifth bar in the chart, has failed the task despite nearly 2000 steps. This pattern offers insights into the efficacy of energy usage in contexts where social abilities are misaligned.

Optimal Social Ability. Fig. 8-(c) and Fig. 8-(d) demonstrate a markable increase in the success rates for teams with ON4 and ON5 social abilities. Specifically, Fig. 8-(c) highlights the advancements made by ON4 teams in both task performance and energy efficiency. It can be seen that the successful rate climbs to 60% but the number of steps average is around 20 steps. It demonstrates a great improvement in collective intelligence. Furthermore, the outcomes for ON5 teams are more impressive, evidenced by an 85% success rate and a reduced NOS for task completion. Teams can find more optimal policies with higher energy efficiency at the team level.

These results indicate that possessing optimal social abilities is crucial for a team in efficiently tackling complex tasks, thereby achieving higher levels of team intelligence. Conversely, a deficiency in social abilities can lead to poor team energy efficiency and potentially deteriorate task performance.

5.2 Teamwork Division Analysis

As discussed in the previous section, the impact of social abilities has been examined at the team level, revealing the great influence it has on teamwork energy efficiency. However, at the agent level, the distribution of energy within a team and the underlying factors that contribute to variations in energy utilization are still lacking. Therefore, our focus shifts to the analysis of how social abilities influence the allocation of workload among team members during the execution of a policy, delving into the mechanisms of teamwork at the agent level.

Based on previous experiments [4] [10] [22], it is reasonable to operate under the presumption that the team size is appropriate, and every agent involved in a task is expected to contribute positively. Hence, it is desirable for teams to learn an effective

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

approach to managing and allocating tasks to each team member. In order to explore the teamwork details, we selected the most successful team from the collective performance of all four teams as an example of the highest group intelligence achieved through training.

Ineffective agents are defined as those agents involved in conflicted actions, as introduced in section 4 and shown in Fig. 7. Our objective is to develop joint policies that minimize the number of ineffective agents. Thus, we track the number of ineffective agents in the Best NS, the Best ON2, the Best ON4, and the Best ON5 teams among 20 pieces of training in completing tasks.

Those results are depicted in Fig. 9 and those are from the fifth bar in Fig. 8-(a), the seventeenth bar in Fig. 8-(b), the nineteenth bar in Fig. 8-(c), and the nineteenth bar in Fig. 8-(d). In Fig. 9, the number of ineffective agents in the above cases are shown by blue, red, purple, and green dashed lines with markers. The graph uses the x-axis to denote the step number, while the y-axis quantifies the number of ineffective agents from the 10-agent team at each given step.

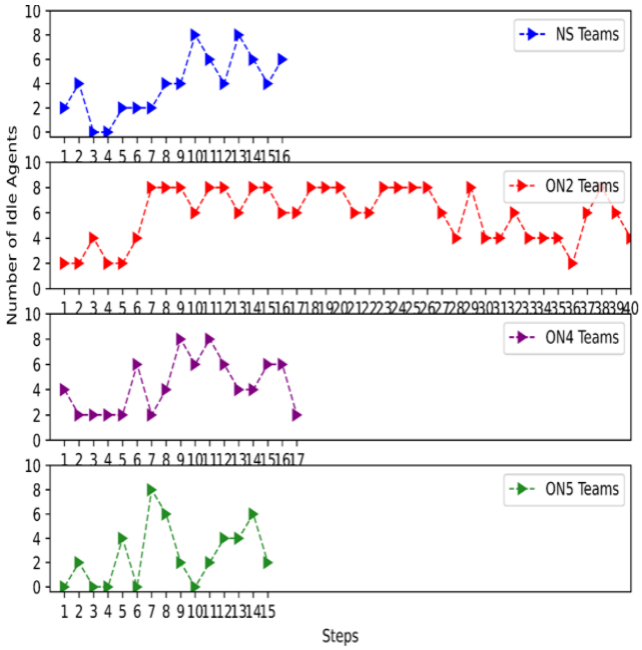


Figure 9: *Teamwork division of different teams.*

Fig. 9-(a) illustrates that the Best NS team completes the task within 16 steps. Nevertheless, this team frequently has over four *ineffective agents* out of ten members. This indicates that relying solely on individual learning makes it hard to facilitate efficient teamwork, leading to energy dissipation via conflicted actions. The integration of social abilities, as shown by the red, purple, and green lines in Fig. 9, exhibits significant pattern variations from the benchmark. This highlights the clear influences of social abilities on the teamwork division.

Insufficient Social Ability. The Best ON2 team exhibits a tendency to take more steps to complete the task, as well as display a higher number of ineffective agents, often exceeding 6 out of 10 team members. Furthermore, the total number of steps exceeds that of other scenarios. The results imply that social abilities stimulate teams to broaden their search within the exploration space. However, inadequate social abilities may hinder the team's capacity to effectively exploit policies, resulting in risks of a lower level of cooperation.

Optimal Social Ability. In contrast, when social abilities are set to be more suitable for this task, the teamwork division is shown to be more reasonable, resulting in fewer required steps and more contributors in the task implementation process. It can be seen that the Best ON4 team is capable of completing the task with 17 steps, which is much fewer than the Best ON2 team. Additionally, the number of active contributors during the task implementation process is higher. The most exemplary performance is observed in the Best ON5 team, requiring only 15 steps with a highly efficient teamwork division, In the Best ON5 team, ineffective agents are reduced to fewer than 4, at times even zero, throughout the task implementation process. This indicates the beneficial impact of suitable social abilities on the efficient allocation of work within a team.

5.3 Energy Utilization Rate

In this section, we conduct a qualitative assessment of social ability's impact on teamwork by examining energy utilization within a team. This metric serves as an indicator of determining the percentage of energy that yields effective results for deepening the understanding of the pattern of energy allocation and consumption and delving into the reasons behind the teams' varying performances.

As mentioned previously, each action "push" taken by each agent is assumed to consume 1 unit of energy. Therefore, we determine the Total Energy Consumption for a single task by multiplying the NOS by its team size (all are set as 10), as shown in equation (13). Considering we have already highlighted the number of ineffective agents in Section 5.2, we now first compute the wasted energy ratio, which is the energy utilized by ineffective agents over the total consumed energy. The energy consumed by ineffective agents, the efforts of ineffective agents, is the aggregated number of ineffective agents at each step throughout the entire task process, as outlined in equation (14). With these calculations, then we can derive the energy utilization rates as shown in equation (15).

$$\text{Total Energy Consumption} = \textit{Number of Steps} \cdot \textit{Team Size} \tag{13}$$

$$\text{Energy Consumed by Ineffective Agents} = \sum_{step=1}^{Number\ of\ Steps} \textit{Number of Ineffective Agent at Current Step} \tag{14}$$

$$\text{Energy Utilization Rate} = 1 - \frac{\textit{Energy consumed by Ineffective Agents}}{\textit{Total Energy Consumption}} \tag{15}$$

Table 5 presents the performance outcomes of the Best NS team, the Best ON2 team, the Best ON4 team, and the Best ON5 team. Notably, the Best NS-Team shows 62 instances of ineffective actions among 160 total unit trials, resulting in an energy utilization rate of 0.6125. However, the introduction of social abilities in the Best ON2 team leads to a substantial increase in ineffective actions, with 234 occurrences, which significantly

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

reduced the energy utilization rate to 0.4150. Conversely, the Best ON4 team demonstrates a significant reduction in ineffective actions to 74 instances, while the Best ON5 team achieved the most efficient performance with only 40 instances of ineffective actions. Those figures translate to higher energy utilization rates, which are 0.5647 and 0.7333, respectively, when compared to the benchmark.

Table 5: *Energy utilization ratio.*

The Best Team	NS	ON2	ON4	ON5
Total Energy Consumption	10*16	10*40	10*17	10*15
Energy Consumed by Ineffective Agents	62	234	74	40
Energy Utilization Ratio	0.6125	0.4150	0.5647	0.7333

These findings highlight that teams lacking social abilities exhibit poorer work division and lower energy utilization rates. In contrast, teams equipped with sufficient social abilities are able to allocate work more reasonably and have fewer inefficient agents involved at each step. This strategic allocation and cooperation lead to superior task performance and greater energy efficiency.

6. CONCLUSION

In multiagent deep reinforcement learning, the abstraction of the learning process makes the team behavior codes hidden behind neural networks. This obscured learning leads to great difficulties in understanding the mechanisms underlying teamwork. Therefore, the progress toward developing teams with high collective intelligence for tackling complex engineering tasks is challenging. To empower the multiagent teams, our research has explored the way of utilizing social learning, a common phenomenon in nature, in MARL to encourage mutual learning. However, the understanding of how social abilities precisely affect teamwork dynamics, training efficacy, and overall performance is limited. Therefore, investigating the role of social abilities is crucial for their strategic and effective incorporation into a team.

This paper has delved into the impact of social abilities on teamwork performance and energy utilization in the context of assembly tasks involving collision avoidance. The findings contribute to the design of more efficient and capable AI robot systems. The experimental findings have led to several key conclusions:

- At the team level, appropriate social abilities are crucial for a team to formulate intelligent strategies that yield efficient energy usage and high-quality teamwork performance. These abilities can enable agent teams to broaden their exploration of optimal policies. On the other hand, inadequate social abilities may hinder teams from fully exploiting the policies due to insufficient social interaction, leading to poor team energy efficiency and worse task performance.
- At the agent level, teams with appropriate social abilities are more capable of allocating work effectively among the team members, resulting in fewer ineffective agents and enhanced collective intelligence. Conversely, teams deficient in these skills may experience inefficiencies in task allocation and reduced energy efficiency.

In conclusion, social abilities are necessary for a team to explore and identify optimal group strategies for a complex task. The correct application and configuration of the social ability are critical for successful training outcomes, particularly in engineering applications where energy efficiency and economic considerations are important. The study has enriched our comprehension of the social mechanisms within learning-based SOSs, suggesting a mapping between desired social information with a targeted task. Designers are advised to thoughtfully integrate proper social skills into teams, carefully evaluating the potential benefits again costs. It is noted that the conclusions presented are based on simulations conducted in this research. Moving forward, we will further deepen our exploration of the impact of social abilities and higher capabilities within more complex scenarios. Additionally, we will explore more generalized cases to bridge the gap between biases in simulation and real-world environments.

ACKNOWLEDGEMENTS

This paper is based on the work supported in part by the Autonomous Ship Consortium (ASC). The authors are grateful for their support.

REFERENCES

[1] U. Othman and E. Yang, "Human–robot collaborations in smart manufacturing environments: review and outlook," *Sensors,* vol. 12, p. 5863, 2023.

[2] N. Gjeldum, A. Aljinovic, C. Zizic and M. Mladineo, "Collaborative robot task allocation on an assembly line using the decision support system," *International Journal of Computer Integrated Manufacturing,* vol. 34, no. 4-5, 2022.

[3] M. K. Habib and Y. Baudoin, "Robot-assisted risky intervention, search, rescue and environmental surveillance," *International Journal of Advanced Robotic Systems,* vol. 7, no. 1, p. 10, 2010.

[4] H. Ji and Y. Jin, "Designing Self-Organizing Systems With Deep Multi-Agent Reinforcement Learning." International Design Engineering Technical Conferences and Computers and Information in Engineering Conference," *American Society of Mechanical Engineers,* p. Vol. 59278, 2019.

[5] C. E. Hrabia, M. Lutzenberger and S. Albayrak, "Towards adaptive multi-robot systems: Self-organization and self-adaptation," *The Knowledge Engineering Review,* vol. 33, p. e16, 2018.

[6] J. Hu and M. P. Wellman, "Multiagent reinforcement learning: theoretical framework and an algorithm," *ICML,* vol. 98, pp. 242-250, 1998.

[7] C. Chen and Y. Jin, "A behavior based approach to cellular self-organizing systems design," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2011.

[8] A. Bandura and R. H. Walters, "Social learning theory," in *Social learning theory*, Prenticd Hall: Englewood cliffs, 1977, p. Vol.1.

[9] C. M. Heyes, "Social learning in animals: categories and mechanisms," *Biological Reveiws,* vol. 69, no. 2, pp. 207-231, 1994.

[10] B. Huang and Y. Jin, "Social learning in self-organizing systems for complex assembly tasks," *Advanced Engineering Informatics,* vol. 57, p. 102109, 2023.

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026

[11] O. Hamid, K. Dautenhahn and C. L. Nehaniv, "Engineering social learning mechanisms for minimalistic multi-agent robots," in *International Conference on Control and Robots (ICCR), IEEE*, 2020.
[12] B. Chen, C. Yu, Q. Diao and R. Liu, "Social or individual learning? An aggregated solution for coordination in multiagent systems," *Journal of Systems Science and Systems Engineering,* vol. 27, pp. 180-200, 2018.
[13] J. Hao, D. Huang, Y. Cai and H.-f. Leung, "The dynamics of reinforcement social learning in networked cooperative multiagent systems," *Engineering Applications of Artificial Intelligence,* vol. 58, pp. 111-122, 2017.
[14] Y. Wang and P. M. Djuric, "Social learning with heterogeneous agents and sequential decision making," *Digital Signal Processing,* vol. 47, pp. 17-24, 2015.
[15] D. Borsa, B. Piot, R. Munos and O. Pietquin, "Observational learning by reinforcement learning," *arXiv preprint,* p. 1706.06617, 2017.
[16] K. K. Ndousse, D. Eck, S. Levine and N. Jaques, "Emergent social learning via multi-agent reinforcement learning," in *International Conference on Machine Learning*, 2021.
[17] G. Chen, "A New Framework for Multi-Agent Reinforcement Learning--Centralized Training and Exploration with Decentralized Execution via Policy Distillation," *arXiv preprint arXiv:1910.09152,* 2019.
[18] L. Buşoniu, R. Babuška and B. D. Schutter, "Multi-agent reinforcement learning: An overview," *Innovations in multi-agent systems and applications-1,* pp. 183-221, 2010.
[19] D. Sam, D. Kudenko and M. Grześ, "An empirical study of potential-based reward shaping and advice in complex, multi-agent systems," *Advances in Complex Systems ,* pp. 251-278, 2011.
[20] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare and A. Graves, "Human-level control through deep reinforcement learning," *nature,* pp. 529-533., 2015.
[21] B. Huang and Y. Jin, "Reward shaping in multiagent reinforcement learning for self-organizing systems in assembly tasks," *Advanced Engineering Informatics,* p. 101800, 2022.
[22] B. Huang and Y. Jin, "Knowledge Transfer in Self-Organizing Systems and Impact of Social Ability," *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference,* vol. 87318, p. V03BT03A004, 2023.
[23] Y. Du, L. Han, M. Fang, J. Liu, T. Dai and D. Tao, "Liir: Learning individual intrinsic reward in multi-agent reinforcement learning," *Advances in Neural Information Processing Systems ,* 2019.
[24] L. Wang, Y. Zhang, Y. Hu, W. Wang, C. Zhang, Y. Gao, J. Hao and T. Lv, "Individual Reward Assisted Multi-Agent Reinforcement Learning," in *International Conference on Machine Learning*, 2022.
[25] V. Klingspor, J. Demiris and M. Kaiser, "Human-robot communication and machine learning," in *Applied Artificial Intelligence*, 1997.
[26] S. Ratchev, and N. Lohse, "Data modelling for web enabled design of modular precision assembly devices," *Assembly Automation ,* vol. 24.1, pp. 63-70, 2004.
[27] P. Shinners, *http://www.pygame.org," 2011.*.
[28] M. Wunder, M. L. Littman and M. Babes, "Classes of multiagent q-learning dynamics with epsilon-greedy exploration," in *In Proceedings of the 27th International Conference on Machine Learning*, 2010.
[29] European Commission, "Energy Efficiency Plan," Brussels, 2011.
[30] G. Carabin, E. Wehrle and R. Vidoni, "A review on energy-saving optimization methods for robotic and automatic systems," *Robotics,* vol. 6, no. 4, p. 39, 2017.
[31] Y. Yan, "To Go or Not to Go: On Energy-Aware and Communication-Aware Robotic Operation," *IEEE Transactions on Control of Network Systems ,* vol. 1, no. 3, pp. 218-231, 2014.
[32] W. Chiang and Y. Jin, "Toward a meta-model of behavioral interaction for designing complex adaptive systems," in *IDETC & CIE* 2011.
[33] X. Wang, and Y. Jin. "Transfer Reinforcement Learning: Feature Transferability in Ship Collision Avoidance." In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 87318, p. V03BT03A071. American Society of Mechanical Engineers, 2023.

Copyright © 2024 by ASME

Downloaded from http://asmedigitalcollection.asme.org/IDETC-CIE/proceedings-pdf/IDETC-CIE2024/88377/V03BT03A002/7403140/v03bt03a002-detc2024-143326.pdf by University of Southern California user on 14 May 2026