Needs-guided Robotic Decision-Making based on Independent Reinforcement Learning

Zhaotie Hao ♠, Bin Guo ♠ [1], Zhuo Sun ♠, Lei Wu ♠, Kaixing Zhao ♣ and Zhiwen Yu ♠
♠ School of Computer Science, Northwestern Polytechnical University, Xi'an, China
♣ School of Software, Northwestern Polytechnical University, Xi'an, China
{haozhaotie, leiwu}@mail.nwpu.edu.cn
{guob, zsun, kaixing.zhao, zhiwenyu}@nwpu.edu.cn

Abstract

In human social interactions, decisions are naturally influenced by both individual needs and the needs of others. However, it remains unclear whether cognitive robots exhibit similar needs-guided decision-making characteristics. In this study, we design a collaborative tracking task to evaluate this phenomenon. Specifically, we develop a needs-guided reinforcement learning framework that enables robots to autonomously learn and shape behavior by considering both their intrinsic needs and those of others. Our experiments highlight that the robots' inherent needs play a more crucial role in decision-making than the needs of others. In essence, our model establishes an interpretable foundation for applications in cognitive robotics.

Keywords: needs-guided; reinforcement learning; robotic decision-making

Introduction

Needs play a pervasive role in human social interactions, providing the foundation for human decision-making. Actually, individuals navigate their actions and decisions in alignment with their internal needs (Leont'ev, 1971). Moreover, humans possess the ability to anticipate the needs of others, enabling them to seize anticipated opportunities or proactively avoid predictable trouble (Zhao & Zhao, 2023). Figure 1 exemplifies a human interaction scenario where the police aim to catch the thief. By predicting the thief's internal needs, the police can adjust their own actions accordingly.

While decision-making guided by both an individual's own needs and the anticipated needs of others, referred to as needs-guided decision-making (Bancerek M, 2023), is prevalent in human interactions, research on needs-guided robotic decision-making remains limited.

Actually, human needs are a form of their desires, and the ability to predict the desires of others is commonly referred to as the theory of mind (Premack & Woodruff, 1978) in cognitive science. A great deal of study has been dedicated to imbuing robots with cognitive capabilities to enhance their decision-making capabilities (Wu S A, 2023; Yuan et al., 2022), exemplified by frameworks such as the Belief-Desire-Intention paradigm (Georgeff, Pell, Pollack, & Tambe, 1998). Yet, these efforts often rely on decisions made by black-box neural networks (Oguntola, Campbell, Stepputtis, & Sycara, 2023; Rabinowitz, Perbet, Song, Zhang, & Eslami, 2018) or Bayesian theory (Wu, Wang, Evans, Tenenbaum, & Parkes, 2021), lacking an exploration of robotic intrinsic motivation. Simultaneously, needs theory is an approach comparable to the theory of mind (Sebastian et al., 2012), highlighting that intrinsic needs drive individuals to take specific actions (Ivancevich, Matteson, & Konopaske, 1990). Although some studies (Yang & Parasuraman, 2023; Sorin, 2009) integrate Maslow's Hierarchy of Needs (Maslow, 1943) into robotic tasks, they often fall short of explicitly verifying the influence of needs on robotic decision-making. Additionally, these approaches tend to overlook the impact of other robots' needs on individual decisions. Consequently, it remains unclear whether needs, both one's own and others needs, can facilitate robotic decision-making tasks.

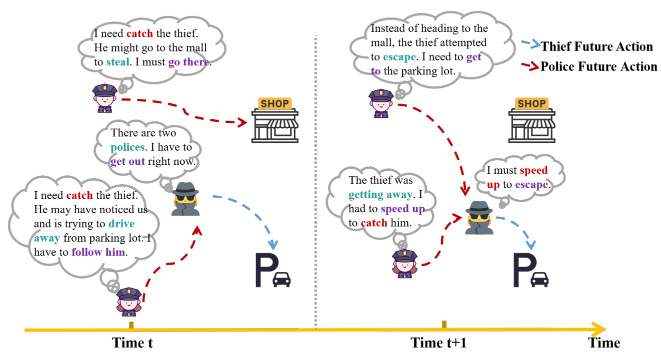


Figure 1: **Example of police catching a thief**. Police officers aim to apprehend the thief promptly, and they predict the thief's needs to infer his future behaviour and guide their own actions. **Left**: The policeman predicts that the thief will commit theft in the mall, while the policewoman predicts that he will drive away from the parking lot. They strategize their actions to apprehend the thief. **Right**: The policeman adjusts his actions, anticipating that the thief will head to the parking lot. Simultaneously, the policewoman observes the thief attempting to escape and accelerates her pursuit.

To address these gaps, we have developed a novel computational framework to investigate how competing models of needs theory can be implemented in robotic decision-making architectures and how they affect robotic decision-making.

In this study, we propose a needs-guided independent reinforcement learning approach designed to address a dynamic task which encompasses both cooperation and confrontation. Specifically, our experimental scenario involves an aerial robot and a ground robot jointly tracking a target. We focus on the integration of robots' own needs and pre-

[1]Corresponding author.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society*.

dictions of target needs for decision-making. Through extensive experiments, our collaborative team showcased the ability to effectively track the target under various constraints. Furthermore, our experiments highlight the general applicability of needs theory within robotic societies. To the best of our knowledge, this is the first model that integrates multiple needs to guide robotic decision-making behaviors.

Related Works

Needs Theory in Psychology

Needs theory is a well-established concept in psychology science, validated through extensive human experience (Ivancevich et al., 1990). Notable theories include Maslow's Hierarchy of needs (Maslow, 1943), the ERG Needs (Alderfer, 1969), Herzberg's Motivation-Hygiene Theory (Hall & Williams, 1986), and McClelland's Needs (McClelland & Boyatzis, 1982).

Maslow's Hierarchy of needs Theory is one of the earliest theories, including physiological, safety, social, esteem, and self-actualisation needs (Maslow, 1943). Maslow believed that there is a hierarchy between these needs, and these needs cannot coexist (Maslow, 1943). However, in reality, humans may experience multiple needs simultaneously. For instance, individuals may strive to fulfill both self-actualisation needs and the basic physiological needs.

Alderfer proposed three fundamental human needs: existence, relatedness, and growth, named as ERG needs (Alderfer, 1969). Specifically, existence needs entail ensuring basic human safety, relatedness needs refer to maintaining interpersonal relationships with others, and growth needs refer to realizing personal aspirations. Unlike Maslow's theory, Alderfer argued for a range rather than a strict hierarchy among these needs. This paper aligns with ERG theory, considering it to be more congruent with human cognition (Acquah, Nsiah, & Antie, 2021), and the experiments conducted are grounded in ERG needs.

Other needs theories concentrate on specific areas. For instance, Herzberg's Theory (Hall & Williams, 1986) focused on job satisfaction, and McClelland's Needs (McClelland & Boyatzis, 1982) proposed developmental needs. These theories are not generally representative.

Needs Theory in Multi-robot Interactions

While needs theory is extensively applied in human interactions, its application in robotic decision-making has received relatively less attention. Actually, once humans have set goals and constraints for robots as their internal needs, robots strive to fulfill these requirements. Additionally, understanding the psychological needs of others enables robots to avoid trouble and develop strategies (Zhao & Zhao, 2023), which is crucial for effective participation in multi-robot collaborative and competitive interactions.

The high efficiency of needs-driven heterogeneous robot collaboration has been demonstrated (Yang & Parasuraman, 2020). However, each robot's needs are treated as independent, which does not align with the existing needs theory. In contrast, (Yang & Parasuraman, 2023) takes a step further by applying Maslow's Hierarchy of Needs (Maslow, 1943) to control a single robot. However, due to the limitations of Maslow's theory, it is difficult for a robot to simultaneously meet multiple needs at any given moment. This scenario is unlikely in reality, where robots must prioritize safety while also fulfilling human tasks. These aforementioned studies do not explicitly explain the relationship between needs and robotic decision-making, including both the robot's own needs and those of others.

Given the scarcity of robotic decision-making using needs theory, this paper attempts to answer two fundamental questions: 1) how ERG-needs theory can be integrated into a robot decision-making architecture, and 2) the relationship between needs and robotic decision-making.

Table 1: Robot settings in the experiment.

Robot Type	X-axis speed (m/s)	Y-axis speed (m/s)	Observable range (m)	Communication range (m)
R_A	-3.5∼3.5	-3.5∼3.5	10	15
R_G	-2.0∼2.0	-2.0∼2.0	2	15
R_T	-2.0∼2.0	-2.0∼2.0	2	–

Method

Task Description

In the considered task, an aerial robot and a ground vehicle robot collaborate to track an intelligent target, which is widely applied in the field of public safety in urban cities (Yu, Han, Chen, Guo, & Yu, 2021). This challenging task involves both intra-team collaboration among robots in a heterogeneous team and team-versus-target confrontation.

Specifically, the heterogeneous team composes of R_A (a UAV) and R_G (a UGV). They collaborate to track a dynamic target R_T (a UGV). Due to the limitation of their observable range, robots cannot obtain global positions when they are beyond their threshold range. Simultaneously, robots must avoid collisions with static obstacles as well as other robots. To synchronize collaboration within the heterogeneous team, R_A and R_G need to maintain a certain communication range to ensure effective interactions. Moreover, each robot makes decisions independently. The overall goal of R_A and R_G is to ensure the target within their observable range, preferably at close range. R_T dynamically adjusts its strategies in real-time to evade tracking by aerial-ground robots. This makes the tracking task more challenging.

Our simulation environment is built on Bullet Physics (Coumans & Bai, 2016), where static square obstacles of varying heights are systematically distributed, presenting potential impediments to the navigation of both aerial and ground-based robotic entities. In each episode, aerial-ground robots start from a free space in the environment randomly, while the target R_T starts around them. R_T adopts adversarial strategies based on real-time situations to evade tracking.

Agent Modeling

For the considered task, we model it as a decentralized partially observable Markov decision process (Dec-POMDP) (Oliehoek & Amato, 2016). The whole process is governed by the tuple $< \mathcal{S}, O_A, O_G, O_T, \mathcal{A}_A, \mathcal{A}_G, \mathcal{A}_T, r_A, r_G, r_T, \mathcal{P} >$, where $\mathcal{S}$, O, $\mathcal{A}$, r, $\mathcal{P}$ denote state space, observation space, action space, reward function, and environment state transition probability, respectively. Notably, the settings of each robot are detailed in Table 1, as we will discuss later. The superscript $t \in \{1, 2, ..., T\}$ denotes the time step. At each time step, robots receive partial observation from the environment and decide actions based on their policies. Then the state is updated from the environment state transition probability $\mathcal{P}$. Meanwhile, each robot receives its immediate rewards from the environment. The robot R_A intends to maximize its expected return $\mathbb{E}_{\pi_A}\left[\sum_{t=1}^{T} r_A^t\right]$ by learning its policy π_A based on model-free independent reinforcement learning. Formally, R_G and R_T have the same objectives as R_A.

Robot observations. Similar to (Young & Tian, 2019), R_A has access to a top-down semantic map along with R_T's orientation. Each pixel of the semantic map corresponds to an object category represented as a one-hot vector, including categories of free space, obstacles, and more. Observations of R_G and R_T are similar to R_A, but have a smaller observable range from their front-view.

Robot decision-making level actions. The ground robot R_G and the target R_T use Mecanum wheels so that they can be controlled by a 2-D continuous vector consisting of linear velocity along the x-axis and y-axis. Similarly, the quadcopter R_A employs a 2-D continuous action space to enable omnidirectional movement.

Communication in the heterogeneous robot team. R_A and R_G in the heterogeneous robot team can freely exchange information, including their observations, individual needs, and predicted target needs, within their communication range. They collaborate by complementing observational information and negotiating to predict needs of the target.

ERG needs and rewards for robots. According to ERG theory (Alderfer, 1969), existence, relatedness, and growth are three fundamental needs of humans, serving as intrinsic factors that drive humans to take specific actions (Caulton, 2012). To verify whether robots with cognitive capabilities also exhibit needs-guided actions as humans do, we introduce the existence need N_{E_i}, the relatedness need N_{R_i}, and the growth need N_{G_i} for the robot i based on ERG theory. Specifically, the ERG needs for heterogeneous team members are

$$N_{E_i} = \sum_{j=1}^{n} \mathbb{1}\left[dist(R_i, obstacles_j) < D_{coll}\right] \tag{1}$$

$$N_{R_i} = \frac{\mathbb{1}\left[dist(R_A, R_G) < D_{comm}\right]}{\max(dist(R_A, R_G), d_{safety})} \tag{2}$$

$$N_{G_i} = \max\left(1 - 2 * \frac{|dist(R_i, R_T) - d^*|}{D_{obs}}, -1\right) \tag{3}$$

where n represents the number of obstacles in the environment. $\mathbb{1}[\cdot]$ represents a conditional judgement, returning 1 if the condition in the brackets is true and 0 otherwise. D_{obs} is the maximum observable range of the robot. d^* is the best tracking distance from the target. Equation (1) represents the existence need N_{E_i} of robot i, indicating its safety performance at a given moment. N_{E_i}'s value increases when the robot moves further away from a risky object (e.g., an obstacle, a robot) beyond its safe range D_{coll}. The relatedness need N_{R_i} of aerial-ground robots ensures that they can communicate within a communication range D_{comm}, which is defined as Equation (2). The growth need could guide robots to perform their assigned tasks to the best of their abilities. The growth need N_{G_i} of aerial-ground robots aims to track the target as closely as feasible, given by Equation (3). When the target is within the robot's field of view (D_{obs}), we set the need as a negative linear correlation between their distances. When the target is outside the robot's D_{obs}, the robot's decision is meaningless (Zhong, Sun, Luo, Yan, & Wang, 2019), and the need is set to -1.

The existence need of the target (N_{E_T}) is similar to that of aerial-ground robots, as referred to in Equation (1). The relatedness need N_{R_T} must ensure that the target remains as far away from being tracked as feasible, and the growth need N_{G_T} aims to get close to the destination to carry out the task. These needs are defined as follows:

$$N_{R_T} = -\sum_{i=1}^{2} \max\left(1 - 2 * \frac{|dist(R_i, R_T) - d^*|}{D_{obs}}, -1\right) \tag{4}$$

$$N_{G_T} = \max\left(1 - 2 * \frac{|dist(Destination, R_T) - d^*|}{D_{obs}}, -1\right) \tag{5}$$

To avoid being tracked, the target requires to maintain a safety distance d_{safety} from aerial-ground robots. Consequently, N_{R_T} is defined as Equation (4). The growth need of the target is to arrive at the destination, defined as Equation (5).

According to ERG theory (Alderfer, 1969), individuals invariably aim to maximize the satisfaction of their needs when making decisions (Machina, 1990). Similarly, reinforcement-learning-based robots aim to maximize their rewards from the environment when making action decisions. Consequently, a correlation can be drawn between ERG needs and rewards, where ERG needs function as intrinsic rewards for the robot, motivating it to satisfy its needs. As the robot garners higher rewards, it experiences increased satisfaction with its needs. Formally, we state the relationship between rewards and ERG needs as follows:

$$\begin{aligned} r_A &= \mu_1 N_{E_A} + \gamma_1 N_{R_A} + \varepsilon_1 N_{G_A} \\ r_G &= \mu_2 N_{E_G} + \gamma_2 N_{R_G} + \varepsilon_2 N_{F_G} \end{aligned} \tag{6}$$

where $\mu_1, \mu_2, \gamma_2, \gamma_2, \varepsilon_1, \varepsilon_2 > 0$ are tunable parameters. For the heterogeneous team, the total reward is $r_{total} = r_A + r_G$. The tracking and counter-tracking process between the target and heterogeneous robots constitutes a zero-sum game, where the target also makes efforts to meet its own needs. Consequently, its reward is described as

$$r_T = -r_{total} + \mu_3 N_{E_T} + \gamma_3 N_{R_T} + \varepsilon_3 N_{G_T} \tag{7}$$

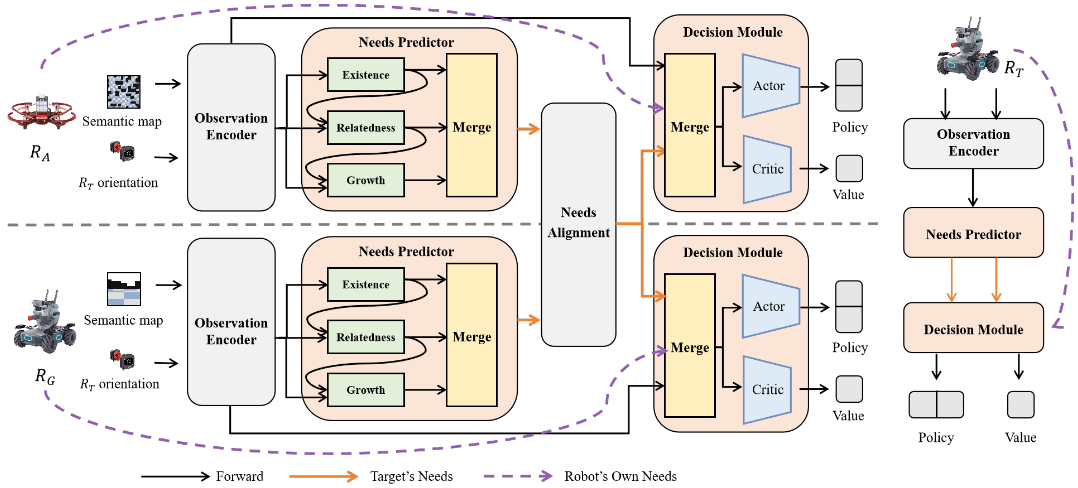


Figure 2: **An overview of our framework**. We establish individualized strategies for aerial and ground robots, respectively. They predict the target needs separately, followed by needs alignment. Each robot make decisions based on both predicted target needs and their own needs.

where $\mu_3, \gamma_3, \varepsilon_3 > 0$ are tunable parameters.

Network Architecture

Figure 2 illustrates an overview of our framework. Each robot has its own unique model, and heterogeneous robots collaborate efficiently through interactions. Our framework consists of an observation encoder, a need prediction module, a value alignment module, and a decision module.

The observation encoder adopts the CNN-LSTM architecture, which comprises two key components. Firstly, a 2-layer CNN (Convolutional Neural Network) (LeCun, Bottou, Bengio, & Haffner, 1998) takes visual observations as input, outputting spatial features. Subsequently, these features merged with orientation observations are fed into a 2-layer MLP (Multi Layer Perceptron) (LeCun, Touresky, Hinton, & Sejnowski, 1988) to encode observations. The encoded features are passed into a LSTM (Long-Short Term Memory) (Graves & Graves, 2012) cell to capture temporal information about the target.

Understanding the needs of others enables robots to avoid trouble and develop strategies (Zhao & Zhao, 2023), crucial for decision-making. Within the heterogeneous team, robots can communicate about each other's needs. However, for a target, aerial-ground robots lack direct communication and can only predict target needs based on encoded observations. According to the ERG theory, human needs are not strictly hierarchical and can be satisfied in any order. Following this principle, we proposed a non-dependent time-series need prediction module. Specifically, each need predictor receives both the encoded observation features and the need predicted by the previous layer of need predictors simultaneously. Similar to ERG theory, we focus on predicting the proportion of the total need that each type of need will account for, rather than specific need values. These proportions sum up to a need share of 1. After predicting the ERG needs separately, these three needs are then merged into a vector $n^*_{i,j}$. Here, $n^*_{i,j}$ denotes the needs of robot j predicted by the robot i.

Due to the huge heterogeneity among aerial-ground robots, their predictions for target needs may differ. Therefore, it is indispensable to align them and approximate the true needs n_T. To align the needs between heterogeneous robots, we introduce a regression task to learn the Needs Alignment Net, parameterized by θ^{VAN}. The loss of this regression task is the Huber loss function:

$$L^N(\theta^{VAN}) = \begin{cases} \frac{1}{2}(n_j - n^*_{i,j})^2 & \left| n_j - n^*_{i,j} \right| \le \delta \\ \delta \left| n_j - n^*_{i,j} \right| - \frac{1}{2}\delta^2 & \text{otherwise} \end{cases} \tag{8}$$

where the tunable parameter $\delta = 1.5$. In accordance with ERG theory, the air-ground robots all make decisions incorporating their own needs, predicted target needs, and encoded observations. The decision model comprises both Actor and Critic components, consisting of a 2-layer MLP with 42 units in the first layer and 100 units in the second layer. The Critic Network estimates the value of the current observed state, while the Actor Network generates actions for the robot. The network architecture of the target is identical to that of the aerial-ground robot, but it directly predicts the needs of the aerial-ground robot without the need for value alignment.

Each robot is trained by the Proximal Policy Optimization (Schulman, Wolski, Dhariwal, Radford, & Klimov, 2017) algorithm with an Adam optimizer (Kingma & Ba, 2014) in an end-to-end manner. The learning rates were 10^{-3} and 10^{-4} for the Actor network and the Critic network respectively. We used a discount factor of 0.9. Its policy loss is:

$$L^P(\theta) = \mathbb{E}_t\left[\min(r_t(\theta)\hat{A}_t, clip(r_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t)\right] \tag{9}$$

where A_t denotes an estimator of the advantage function at timestep t. The term $r_t(\theta)$ denotes the probability ratio between the current stochastic policy and the old stochastic policy with which an robot collected the experience to learn from. The function $clip()$ establishes a bound for the probability ratio term $r_t(\theta)$ within the interval $[1-\varepsilon, 1+\varepsilon]$. In our study, $\varepsilon = 0.2$. The total loss of overall network is:

$$L(\theta) = L^N(\theta^{VAN}) + L^P(\theta) \tag{10}$$

Results and Discussion

The Benefit of Needs

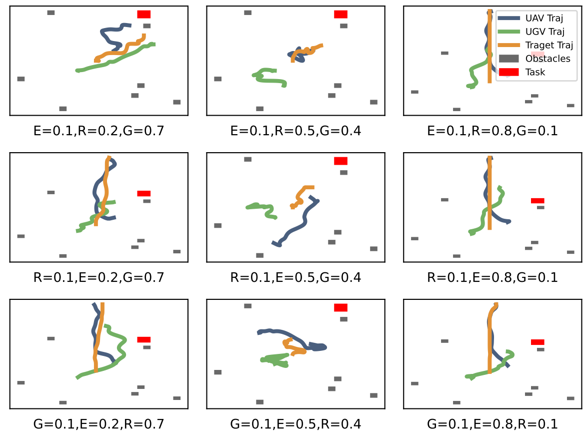
For a visual illustration of the impact of needs on robotic decision-making, we plotted the 2D trajectories of robots during an episode in testing environments in Figure 3. In this visual analysis, we maintained one ERG need proportion constant while varying the proportions of the other two ERG needs. Importantly, to compare the effects of different needs on robotic decision-making, we evaluate the trajectory trends of robots under various needs while maintaining a consistent parameter design and initial values for each experiment.

As depicted in Figure 3a, the trajectories of the robots change significantly when fixing one need and altering the share of the other two needs. This demonstrates that different percentages of others' ERG needs can exert distinct effects on robotic decision-making. It is crucial to highlight that while the contribution of the target needs to the decision-making of the aerial-ground robot is fixed, the true needs of the target do not remain constant. Instead, they change dynamically, leading to a large deviation between the true needs of the target and the target needs used as decision-making input for the aerial-ground robot. This dynamic nature of the robot's own needs may potentially lead to task failure, as observed in the case of subgraph $R = 0.1, E = 0.5, G = 0.4$.

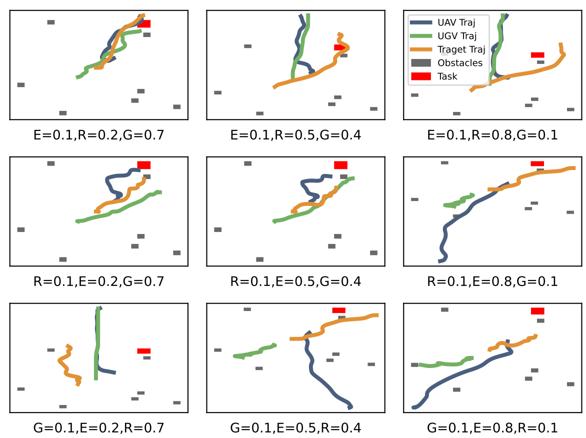
As illustrated in Figure 3b, the robot's own needs also have an effect similar to that observed in Figure 3a, exerting a more substantial impact on robotic decision-making. When the robot's own needs are fixed, it struggles to dynamically assess its own needs in real-time, resulting in occasional significant deviations in its trajectory. In addition, it is intriguing to note that when the robot itself has a disproportionately high share of a particular need, exemplified by the subgraph $E = 0.1, R = 0.8, G = 0.1$, the robot will tend to prioritize satisfying this need more. Since the robot's own need is fixed in these scenarios, it will never be satisfied, leading to the offset trajectories that sometimes occur.

Despite Figure 3 revealing that when robots exhibit diverse needs, their decisions undergo significant variations, it is crucial to emphasize that fixing a particular need does not guarantee effective tracking by heterogeneous robots. Self-needs and the needs of others, along with interactions within each need and observations of robots, dynamically change over time. This dynamic nature makes it challenging for robots to achieve effective tracking.

Furthermore, it is intriguing to note that our trajectories align seamlessly with ERG theory. There is no defined hierarchical order of ERG needs, and lower-level needs are not a prerequisite for satisfying higher-level needs. For instance, in the $E = 0.1, R = 0.2, G = 0.7$ case in Figure 3b, the aerial robot would prioritise the satisfaction of the growth need over the existence need, and then it would move fast to its destination even if there were obstacles.



(a) Fixed Predicted Others Needs

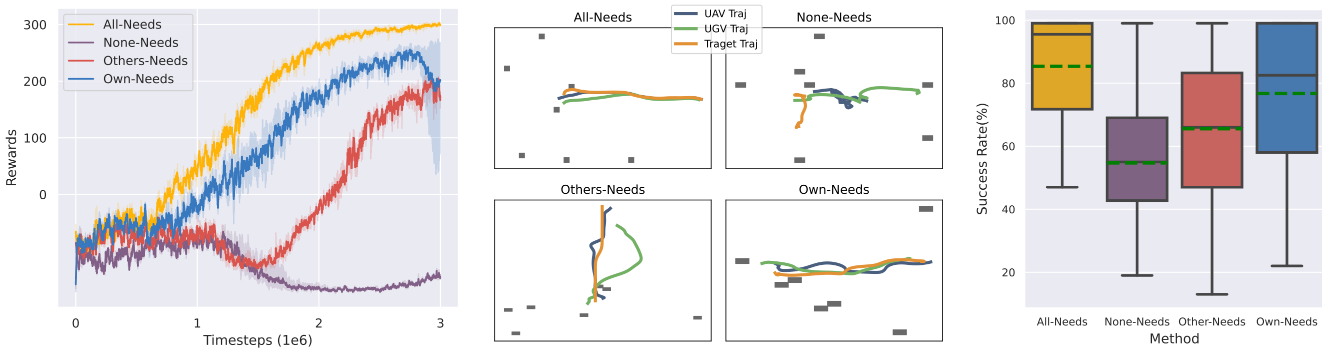


(b) Fixed Own Needs

Figure 3: **Qualitative experiments**. Initially, we trained the models for robots individually at various ERG need proportions. Subsequently, as depicted in the individual subfigures, we conducted separate tests based on the same ERG needs for each trained model. Notably, we tested the robots' own needs and others' needs independently, providing a thorough evaluation of their impact on the robots' action decisions.

Which Needs Taking Priority

To assess the impact of needs on robotic decision-making, as illustrated in Figure 4, we compared the cumulative rewards of the heterogeneous team across several approaches, namely All-Needs (utilizing both own needs and target needs), None-Needs (not utilizing needs), Own-Needs (utilizing only own needs), and Others-Needs (utilizing only target needs). More precisely, Figure 4a graphically depicts the tracking performance of these methods. All-Needs method outperforms other baselines and converges rapidly. In contrast, the None-Needs method performs poorly and fails to converge. This demonstrates the significance of the need for aerial-ground



(a) Accumulated Rewards (b) Trajectory (c) Tracking Success Rate

Figure 4: **Ablation experiments**. **(a)**: accumulated rewards in training environments with different forms of needs. **(b)**: robot trajectories for different needs methods in a test environment. **(c)**: tracking success rate of different formal needs methods in a test environment. The green dashed line indicate the average success rate of each method over 100 episodes, and the solid black line indicates the median.

collaboration. Comparatively, the Others-Needs method's performance rises slowly and is sensitive to random seed. The own-needs method converges rapidly, but it falls short of the All-Needs method in terms of final performance. This demonstrates that the robot's decision-making prioritizes its own needs over those of others.

While both the Own-Needs method and the Others-Needs method can eventually learn the target's strategy and achieve convergence based on observations of the target, they are relatively time-consuming. In contrast, the All-Needs method integrates both its own needs and the target's needs, capturing the intrinsic motivation of the target and swiftly yielding excellent tracking results.

To visually demonstrate the tracking performance of these methods, we plot the trajectory of the robot during the tracking process in Figure 4b. The All-Needs method can quickly track the target, while other methods exhibit poor tracking performance. Robots without ERG needs struggle to capture the motivation of the target, leading to the target and the aerial-ground robot moving in different directions. This once again proves the positive effect of needs on collaborative robotic decision-making. Additionally, the Others Needs method, which relies only on the target's need for decisions, is not effective enough for the ground robot to capture their intrinsic motivations due to its smaller field of view. In contrast, the own-needs method can track the target more effectively, implying that robots prioritize their own needs over those of others when making decisions.

We also conducted additional tests to evaluate the tracking success rate of the aerial-ground team across 100 episodes, as depicted in Figure 4c. In each episode comprising 100 timesteps, tracking is deemed unsuccessful when the target exits the field of view of the aerial-ground team. As seen from Figure 4c, the mean value of the Own-Needs method surpasses that of the Other-Needs method, underscoring the significance of prioritizing one's own needs. While each method achieves a high success rate due to the robot's observations, it is undeniable that the integration of needs plays a pivotal role in facilitating robotic decision-making, especially in scenarios that demand heightened tracking performance.

General Discussion

In this study, we demonstrate the effectiveness of an ERG needs-guided reinforcement learning approach in enhancing robotic decision-making and fostering efficient collaboration among heterogeneous robots in a dynamic tracking task. Our results demonstrate that the robot's own needs play a significant role in shaping its decision-making process, while the needs of other robots also contribute to the overall decision-making of the robot. Moreover, our simulation experiments further validate the general adaptability of the ERG theory within the robotic society, underscoring the time-independent nature of various needs.

In terms of the future of cognitive robotics, our experiments demonstrate that configuring the robot with appropriate ERG needs proportions enables effective guidance for successfully completing tasks assigned by humans. Furthermore, our independent needs prediction module serves as a means for comprehending robots behaviour and providing explanations for robots' behavioural decisions. Our findings align with the ERG theory, suggesting that our framework can serve as a valuable tool for validating needs theories through computer simulations. In summary, our approach lays the foundation for further exploration of needs theory and its applications in robotics.

Despite the key insights gained from our study, there remain promising avenues for future research: (1) Additional investigation into the combinations of ERG needs leading to specific behavioral actions is required. (2) Explore integrating diverse needs for producing specific collaborative behaviors among multiple robots. Hence, future research will delve into how to integrate various needs to generate specific collaborative behaviours among multiple robots.

Acknowledgments

This work was supported by National Science Fund of China for Distinguished Young Scholars 62025205; in part by the National Natural Science Foundation of China under Grant 62032020, and Grant 62102317. Bin Guo is the corresponding author of this paper.

References

Acquah, A., Nsiah, T. K., & Antie, E. N. A. (2021). Literature review on theories of motivation. *EPRA International Journal of Economic and Business Review*, *9*(5), 25–29.

Alderfer, C. P. (1969). An empirical test of a new theory of human needs. *Organizational behavior and human performance*, *4*(2), 142–175.

Bancerek M, Z. J. (2023). The emergence of coordinative dialogue–pragmatic context in multi-agent communication. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 45, pp. 686–692).

Caulton, J. R. (2012). The development and use of the theory of erg: A literature review. *Emerging Leadership Journeys*, *5*(1), 2–8.

Coumans, E., & Bai, Y. (2016). *Pybullet, a python module for physics simulation for games, robotics and machine learning.* https://pybullet.org/wordpress/.

Georgeff, M., Pell, B., Pollack, M., & Tambe, M. (1998). The belief-desire-intention model of agency. In *Intelligent agents v: Agents theories, architectures, and languages: 5th international workshop.* Springer.

Graves, A., & Graves, A. (2012). Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, 37–45.

Hall, J., & Williams, M. S. (1986). *Work motivation inventory. the woodlands, texas: Telemetrics international.* Inc.

Ivancevich, J. M., Matteson, M. T., & Konopaske, R. (1990). *Organizational behavior and management*. McGraw-Hill Education: Prentice-Hall.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

LeCun, Y., Touresky, D., Hinton, G., & Sejnowski, T. (1988). A theoretical framework for back-propagation. In *Proceedings of the 1988 connectionist models summer school* (Vol. 1, pp. 21–28).

Leont'ev, A. (1971). Potrebnosti, motivy i jemocii. *M.: Izd-vo Mosk. un-ta*, *187*, 8.

Machina, M. J. (1990). Expected utility hypothesis. In *Utility and probability* (pp. 79–95). Springer.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*, 370-396.

McClelland, D. C., & Boyatzis, R. E. (1982). Leadership motive pattern and long-term success in management. *Journal of Applied psychology*, *67*(6), 737.

Oguntola, I., Campbell, J., Stepputtis, S., & Sycara, K. (2023). Theory of mind as intrinsic motivation for multi-agent reinforcement learning. *arXiv preprint arXiv:2307.01158*.

Oliehoek, F. A., & Amato, C. (2016). *A concise introduction to decentralized pomdps* (Vol. 1). Berlin: Springer.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, *1*(4), 515–526.

Rabinowitz, N., Perbet, F., Song, F., Zhang, C., & Eslami, S. A. (2018). Machine theory of mind. In *International conference on machine learning* (pp. 4218–4227).

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Sebastian, C. L., Fontaine, N. M., Bird, G., Blakemore, S.-J., De Brito, S. A., McCrory, E. J., & Viding, E. (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social cognitive and affective neuroscience*, *7*(1), 53–63.

Sorin, M. (2009). A model of cooperative agent based on imitation and maslow's pyramid of needs. In *International joint conference on neural networks* (pp. 1229–1236).

Wu, S. A., Wang, R. E., Evans, J. A., Tenenbaum, J. B., & Parkes, D. C. (2021). Too many cooks: Bayesian inference for coordinating multi-agent collaboration. *Topics in Cognitive Science*, *13*(2), 414–432.

Wu S A, G. T., Sridhar S. (2023). A computational model of responsibility judgments from counterfactual simulations and intention inferences. In *Proceedings of the annual meeting of the cognitive science society* (pp. 3375–3382).

Yang, Q., & Parasuraman, R. (2020). Needs-driven heterogeneous multi-robot cooperation in rescue missions. In *2020 ieee international symposium on safety, security, and rescue robotics (ssrr)* (pp. 252–259).

Yang, Q., & Parasuraman, R. (2023). A strategy-oriented bayesian soft actor-critic model. *Procedia Computer Science*, *220*, 561–566.

Young, K., & Tian, T. (2019). Minatar: An atari-inspired testbed for thorough and reproducible reinforcement learning experiments. *arXiv preprint arXiv:1903.03176*.

Yu, Z., Han, L., Chen, C., Guo, W., & Yu, Z. (2021). Object tracking by the least spatiotemporal searches. *IEEE Internet of Things Journal*(16).

Yuan, L., Gao, X., Zheng, Z., Edmonds, M., Wu, Y. N., Lu, H., … Zhu, S.-C. (2022). In situ bidirectional human-robot value alignment. *Science robotics*, *7*(68), eabm4183.

Zhao, Z., & Zhao, F. (2023). A brain-inspired theory of mind spiking neural network improves multi-agent cooperation and competition. *Patterns*.

Zhong, F., Sun, P., Luo, W., Yan, T., & Wang, Y. (2019). Ad-vat+: An asymmetric dueling mechanism for learning and understanding visual active tracking. *IEEE transactions on pattern analysis and machine intelligence*, *43*(5), 1467–1482.