

# MHRC: Closed-loop Decentralized Multi-Heterogeneous Robot Collaboration with Large Language Models

Wenhao Yu<sup>1</sup>, Jie Peng<sup>2</sup>, Yueliang Ying<sup>3</sup>, Sai Li<sup>4</sup>, Jianmin Ji<sup>4,\*</sup> and Yanyong Zhang<sup>2</sup>

**Abstract**—The integration of large language models (LLMs) with robotics has significantly advanced robots’ abilities in perception, cognition, and task planning. The use of natural language interfaces offers a unified approach for expressing the capability differences of heterogeneous robots, facilitating communication between them, and enabling seamless task allocation and collaboration. Currently, the utilization of LLMs to achieve decentralized multi-heterogeneous robot collaborative tasks remains an under-explored area of research. In this paper, we introduce a novel framework that utilizes LLMs to achieve decentralized collaboration among multiple heterogeneous robots. Our framework supports three robot categories—mobile robots, manipulation robots, and mobile manipulation robots—working together to complete tasks such as exploration, transportation, and organization. We developed a rich set of textual feedback mechanisms and chain-of-thought (CoT) prompts to enhance task planning efficiency and overall system performance. The mobile manipulation robot can adjust its base position flexibly, ensuring optimal conditions for grasping tasks. The manipulation robot can comprehend task requirements, seek assistance when necessary, and handle objects appropriately. Meanwhile, the mobile robot can explore the environment extensively, map object locations, and communicate this information to the mobile manipulation robot, thus improving task execution efficiency. We evaluated the framework using PyBullet, creating scenarios with three different room layouts and three distinct operational tasks. We tested various LLM models and conducted ablation studies to assess the contributions of different modules. The experimental results confirm the effectiveness and necessity of our proposed framework.

## I. INTRODUCTION

The complexity of real-world tasks poses significant challenges in designing a single robot capable of handling all aspects of task execution. This has driven a growing body of research focused on multi-heterogeneous robot collaboration [1]. These systems often consist of robots with varying capabilities, such as manipulation and navigation, working collaboratively to achieve common objectives [2]. In this task context, traditional approaches often rely on explicit programming and centralized control, which tend to be inflexible and inefficient in managing the complexities of real-world scenarios [3].

Recent developments in LLMs offer a promising avenue for enhancing robot collaboration through natural language

<sup>1</sup> Institute of Advanced Technology, University of Science and Technology of China (USTC), Hefei 230026, China  
wenhaoyu@mail.ustc.edu.cn

<sup>2</sup> School of Artificial Intelligence and Data Science, USTC, 230026, China

<sup>3</sup> The University of North Carolina at Chapel Hill, USA

<sup>4</sup> School of Computer Science and Technology, USTC, 230026, China

\* Corresponding author: jianmin@ustc.edu.cn

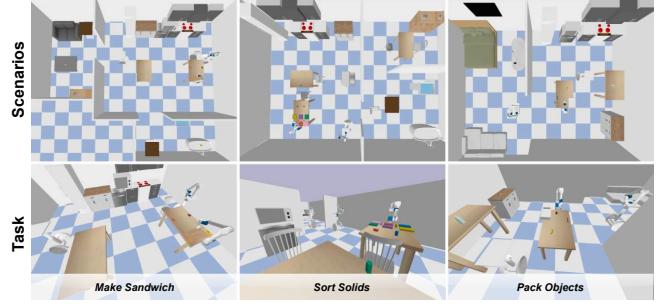


Fig. 1. This figure depicts the experimental scenarios and tasks involved in our research. The furniture types and layouts differ across various settings, with distinct room configurations, such as kitchens, bathrooms, and bedrooms. The task design is inspired by RoCo [5], encompassing activities such as sandwich making, sorting solid objects, and packing items.

understanding and generation. Natural language provides a unified interface for robot collaboration, enabling robots to better understand each other’s actions and communicate more conveniently. Recent advancements in LLMs have demonstrated remarkable capabilities in natural language understanding, dialogue generation, extensive world knowledge, and complex reasoning. Leveraging these strengths offers new opportunities to address challenges in decentralized multi-heterogeneous robot collaboration [4], [5].

Therefore, we introduce MHRC, a novel framework that harnesses the power of LLMs to enable decentralized multi-heterogeneous robot decision-making and planning. We model the collaborative task as a decentralized partially observable Markov decision process (DEC-POMDP) [6], [7], which allows each robot to make decisions based on its local observations and communications with other robots. We meticulously designed prompt templates for the tasks, leveraging in-context learning to tap into the extensive world knowledge of LLMs. Furthermore, we employed the CoT [8] framework to enhance the LLMs’ capabilities in processing long sequences and performing complex reasoning.

Our proposed framework consists of three key modules: **Observation Module**: This module gathers all necessary state information for task and motion planning. It includes a scene graph representing structured information about the environment, messages exchanged between robots, and individual robot status information. The scene graph is dynamically updated to reflect changes in the environment, such as the positions of objects and the states of manipulable furniture. The scene graph updates are based on the robot’s local observations and the received messages. Communica-

tion between robots is handled through structured natural language prompts, allowing them to share information and request assistance effectively. **Memory Module:** To handle long-horizon tasks and maintain context across multiple interactions, the memory module records feedback history, received message history, and action history. By marking recent feedback and messages with special tags, the module ensures that the most relevant information influences the robots' decision-making processes. **Planning Module:** Utilizing LLMs, each robot independently generates actions by selecting from a predefined list, guided by the CoT framework. After executing an action, the robot receives feedback and updates its memory, allowing it to adjust its plan accordingly. This iterative process enables robots to adapt to new information and collaborate effectively to complete complex tasks.

We validate our framework through experiments conducted in simulated environments created using PyBullet [9]. The environments feature various room layouts, furniture arrangements, and object distributions to simulate realistic scenarios. We design three distinct tasks—Pack Objects, Sort Solids, and Make Sandwich—to assess the robots' capabilities in basic manipulation, color matching, and sequential stacking, respectively. Each task requires collaboration among three types of robots: a mobile robot responsible for exploration, a mobile manipulation robot that opens furniture and transports objects, and a manipulation robot that performs tabletop operations.

Our evaluation metrics include success rate, partial success rate, average temporal steps, and average action steps. The results demonstrate that our framework enables efficient and effective collaboration among heterogeneous robots, achieving high success rates across all tasks. Notably, the use of LLMs allows for flexible and decentralized control, as each robot can autonomously make decisions based on its observations and communications, without relying on a centralized controller.

The main contributions of our work are summarised as follows:

- The paper introduces MHRC, a novel framework that leverages LLMs for decentralized collaboration among heterogeneous multi-robot systems.
- Leveraging a diverse set of feedback, we developed a tailored replanning mechanism for different types of robots. The mobile manipulation robot can dynamically adjust its base position to optimize grasping conditions, while the manipulation robot demonstrates an enhanced understanding of task requirements, ensuring accurate and reliable execution of actions.
- We have provided a PyBullet-based simulation benchmark, which encompasses three distinct room layout scenarios and three diverse operational tasks.
- We provide comprehensive evaluations demonstrating that our method achieves high success rates and efficient collaboration in complex tasks.

## II. RELATED WORK

### A. *LLM for Robotics*

Integrating LLMs into robotic task planning has led to significant advancements in skill learning, control performance, collaboration, human-robot interaction, and navigation [10]–[15]. LLMs are increasingly used to understand high-level instructions and generate low-level actions that robots can execute directly from textual inputs [14], [16], [17]. Say-Can [14] combines pre-trained skills with value functions to select appropriate actions for embodied tasks. Code as Policies (CaP) [17] uses structured code generation for robot control, while VoxPoser [18] employs 3D Value Maps and LLM-generated code for manipulation tasks, with re-planning capabilities in case of failure. Inner Monologue [19] improves decision-making by using natural language feedback to guide task planning, and ProgPrompt [16] enables plan generation in diverse environments. Recent methods decompose high-level instructions into executable steps using pre-trained LLMs, often in a zero-shot manner [12], [14]. Huang et al. [12] use GPT-3 [20] and Codex [21] to generate action plans for embodied agents, translating each step into an executable action with the help of the Sentence-RoBERTa model [22], [23]. The methods discussed above highlight the substantial potential of LLMs in advancing robot task planning. Our work is to use LLMs to build a framework that can realize multi-heterogeneous robot collaboration tasks.

### B. *Multi-agent Collaboration, Communication, and Motion Planning*

The study of multi-agent systems (MAS) has a rich history [24], [25]. MAS can be classified into two types: homogeneous, where agents share similar characteristics, and heterogeneous, where agents have diverse capabilities [26].

1) *Homogeneous Agents*: Sampling-based methods and their various algorithmic improvements have been proposed are a widely used approach in Homogeneous Agents [27], [28]. Li et al. [29] tackled the optimal synchronization in homogeneous multi-agent systems using an actor-critic neural network and least squares to approximate the control policy and value function. Zhang et al. [30] use a non-policy reinforcement learning algorithm using a single critic neural network to compute each agent's optimal control policy. CoELA [31] leverages LLMs to establish a collaborative framework for homogeneous agents, facilitating cooperative exploration, transport, and communication between the two agents. Guo et al. [32] places greater emphasis on the organizational framework of LLM agents, investigating how to harness the potential of LLMs to develop more effective collaboration strategies.

2) *Heterogeneous Agents*: Compared to homogeneous agents, heterogeneous agents present different challenges, including the heterogeneity of agents, limited view of the environment, and the dynamicity of the multi-agent system (MAS) or environment [33]. Several heterogeneous robot combinations have already been proposed, including aerial-ground collaboration, main-picket collaboration, and

humanoid-quadruped collaboration [34]–[37]. Liu et al. [4] applied LLMs to ad hoc and original heterogeneous robots to generate reasonable collaboration strategies. Zhao et al. [38] developed coordination strategies for robots operating with asymmetric information and varying levels of influence. Haldane et al. [36] address joint locomotion and perception tasks using legged robots of different sizes and capabilities.

### C. LLM with closed-loop feedback for robot replanning

TREE-PLANNER [39] optimizes the closed-loop feedback process through plan sampling, action tree construction, and grounded decision-making. Huang et al. [40] presented a general formulation of Inner Monologue that combines different sources of environmental feedback with methods fusing LLM planning with robotic control policies. Hi-CRISP [41] allows robots to identify and correct errors at each step of task execution. The re-planning for closed-loop feedback in COME-robot systems [42] for mobile robots has been primarily focused on the robotic arm, without fully exploiting the inherent mobility of the robot. This oversight limits the robot’s effectiveness in tasks such as grasping and transporting objects. Compared to these methods, MHRC emphasizes decision-making and planning tasks for multiple heterogeneous robots by integrating closed-loop feedback specifically associated with robotic capabilities with LLMs.

## III. METHOD

### A. Preliminary Information

We formulate this problem as a decentralized partially observable Markov decision process (DEC-POMDP) to facilitate collaborative tasks among heterogeneous multi-robots. Specifically, a DEC-POMDP is defined as  $(n, \mathcal{S}, \{\mathcal{A}_i\}, \{\mathcal{O}_i\}, \mathcal{P}, \mathcal{T}, \mathcal{R}, \gamma, \mathcal{H})$ , where  $n$  is the number of heterogeneous robots in the task;  $\mathcal{S}$  denotes a finite set of states;  $\mathcal{A}_i = A_i^K \times A_i^C$  represents the action set for robot  $i$ , including a finite action set  $A_i^K$  determined by the robot’s individual capabilities and a communication action  $A_i^C$  used to exchange messages with other robots;  $\mathcal{O}_i = O_i^K \times O_i^C$  signifies the observation set for robot  $i$ , encompassing both the observation set  $O_i^K$  derived from the robot’s perceptions and the message set  $O_i^C$  received from other robots;  $\mathcal{P}(s'|s, a)$  and  $\mathcal{T}(o|s')$  are the state transition probabilities and conditional observation probabilities respectively;  $\mathcal{R}$  is the reward function;  $\gamma$  is the discount factor in  $(0, 1)$ ;  $\mathcal{H}$  is a finite planning horizon.

It is important to emphasize that our research centers on optimizing the utilization of LLMs for executing heterogeneous multi-robot collaborative tasks, which fall under the domain of robot task and motion planning(TAMP). Consequently, our approach incorporates a description function  $f$  that translates the semantics and observations related to the robotic tasks into natural language prompts  $l_i^t = f_i(o_i^t), o_i^t \in \mathcal{O}_i$  and does not directly introduce detection-related algorithms within our framework.

### B. Multi-Heterogeneous Robot Collaboration with LLMs

1) *Observation Module*: This module encompasses all the state information necessary for robot tasks and motion planning. It mainly includes the following three parts:

**Scene Graph**: This part presents structured information about the environment. The scene graph encompasses the positions and orientations of key furniture items. For each piece of furniture, at least one navigation target point (robot’s target pose) is defined for “navigate” action. This work is similar to [43]–[45]. In addition to this static information, the scene graph incorporates dynamic updates, including the open/closed status for furniture that requires manipulation and the integration of newly discovered objects, along with updates to their positions and orientations.

**Messages Information**: To achieve effective collaboration in completing tasks, the robots must share information and request assistance as needed. This coordination is enabled through a structured communication process. The experiment begins with the manipulation robot (Bob) initiating two types of requests to the mobile manipulation robot (Alice): one for exploring the environment to locate task-specific objects, and another for transporting objects on the table that are beyond Bob’s operational range. Alice will delegate certain exploration tasks to the mobile robot (David) based on her specific mission requirements. Upon locating the necessary objects, David will relay their positions back to Alice. Notably, all communication occurs via prompts, effectively guiding the LLMs to generate responses autonomously.

**Robot Information**: This section includes the different capabilities of the robot and the corresponding status information. For robots equipped with navigation capabilities, the status includes the robot’s pose information; for robots with manipulation capabilities, the status encompasses the gripper’s open or closed state, the name of the grasped object, and the maximum grasping range.

2) *Memory Module*: In the context of our work, the tasks are designed as long sequences. As the LLMs are queried multiple times, they may lose track of prior tasks, making it essential to retain past decision contexts. The memory module is designed to capture three key components:

**feedback history**: This section primarily stores limited feedback generated from the robot’s past interactions with the environment. We have designed various types of textual feedback for our tasks, covering both success and failure cases. Examples include feedback on the successful execution of atomic actions, updates on the current task status, failure to navigate near the target point, inability to grasp an object due to the suction gripper being too distant, failure to meet task constraints, and failure to satisfy task requirements, among others. For a more comprehensive descriptions of the feedback, please refer to Tab. I.

**received message history**: This part primarily archives messages previously received by the robot from other robots. For detailed information regarding message types, please refer to III-B.1, and for message formats, please refer to Fig. 3

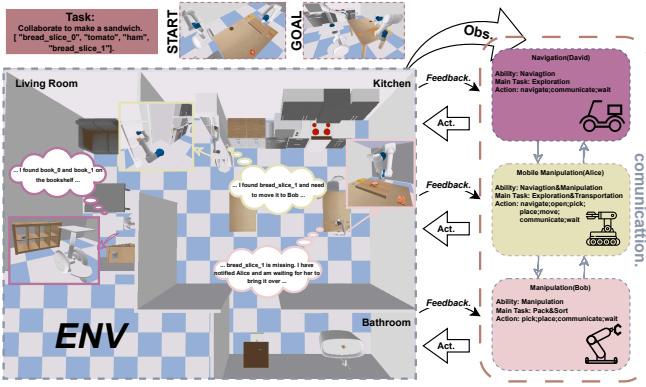


Fig. 2. The figure presents the overall workflow using a representative example. The “Start” and “Goal” denote the initial and target states of a task, respectively. Each robot autonomously makes decisions, plans, and executes atomic actions based on its local observations and communications received from other robots. The robots continuously replan their actions in response to environmental feedback. Through coordinated collaboration, multiple heterogeneous robots work together to accomplish complex, long-sequence tasks.

**action history:** This component primarily archives the sequences of decision-making actions previously executed by the robot, with both types and formats fully adhering to the designed action set III-B.3.

In addition, we propose to mark the most recent feedback and received messages with additional tags, as the latest developments often have a more significant impact on the strategy. This approach ensures that LLMs can prioritize recent information while maintaining a comprehensive understanding of the overall context.

**3) Planning Module:** We employ LLMs to simulate role-playing for various types of robots, guiding task planning through a CoT framework. This process involves selecting and executing one action from a predefined list at a time. After each action, feedback is obtained, and based on this feedback, along with historical data, the robot autonomously determines the next subtask or adjusts the plan. This multi-turn interaction enables robots to collaborate effectively in completing long-horizon tasks.

In mobile manipulation scenarios, identifying an optimal base position is critical, as it significantly influences the success of grasping actions [46]. To address this, we propose a strategy that combines the “*navigate*” and “*move*” actions to accomplish the task. For instance, when grasping an apple from a table, the robot first selects a navigation target from the scene graph and moves near the table. It then evaluates the validity of the robotic arm’s initial and target configurations relative to the apple’s position and attempts to grasp the object. If the attempt fails, the robot selects alternative target points closer to the apple, navigates to those positions, and reattempts the action. Should the task remain incomplete, the robot refines its position using the “*move*” action, adjusting its base location according to the relative x and y coordinates between the base and the apple.

For tabletop manipulation tasks, it is essential to evaluate the success of pick-and-place actions and ensure that the

## Algorithm 1 Multi-Heterogeneous Robot Collaboration

**Require:** heterogeneous robots  $r_0, \dots, r_i, \dots, r_n$ , planning horizon  $\mathcal{H}$

**Require:** memory buffer  $b_i \subset \mathcal{B}$ , prompt policy  $\pi_i \in \Pi$

- 1:  $\mathcal{O} \leftarrow \text{env.reset}()$
- 2: **for** step  $t = 0$  **to**  $\mathcal{H}$  **do**
- 3:     **for** robot  $i = 1$  **to**  $n$  **do**
- 4:          $a_i^t = \pi_i(b_i, f_i(o_i^t))$      ▷ atomic action  $a_i^t \in \mathcal{A}_i$
- 5:          $f_i^t, m_i^t = \text{env.step}(a_i^t)$      ▷ limited feedback  $f_i^t$  from env after  $a_i^t$  is executed and message  $m_i^t$  received from other robots  $r_{n \neq i}$
- 6:          $m_i \leftarrow m_i \cup \{a_i^t, f_i^t, m_i^t\}$
- 7:     **end for**
- 8:      $\mathcal{B} \leftarrow \mathcal{B} \cup b_i$
- 9: **end for**

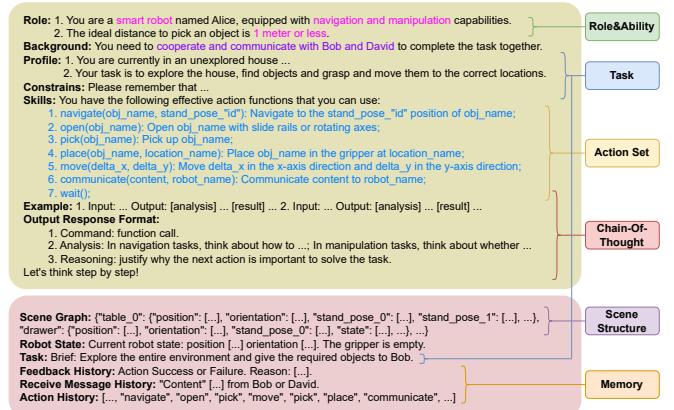


Fig. 3. The key prompts for the work are divided into two components: the system prompt (top), which remains constant throughout the task, while the user prompt (bottom) dynamically evolves in response to the progression of the task. The figure uses a mobile manipulation robot as an example to illustrate prompt design. For prompt design of other robots and more detailed content, please refer to the appendix V-A.

outcomes align with the task’s overall objectives following the execution of the planned motions [41].

## IV. EXPERIMENTS

### A. Environments, Tasks, and Metric

In this study, we constructed several experimental environments using the BestMan [47], [48] simulation platform and developed additional functionalities to support the specific tasks of our research. Each environment features a variety of room types and layouts, diverse furniture types and arrangements, and an array of different objects. The first row of Fig. 1 illustrates the environments corresponding to these three different scenarios, respectively. For the robot navigation algorithm, we use the A\* [49] algorithm for global path planning without considering local obstacle avoidance [50]–[52], as it is not the focus of our task. For the robotic arm planning algorithm, we use the sample-based BIT\* [53] algorithm, considering path planning in a static state.

We also designed the following three different tasks:

TABLE I  
FEEDBACK DESCRIPTION

Type	Description
Feedback on successful actions	
Navigation Success	Provide feedback confirming successful arrival at the designated target point. Additionally, for furniture items such as tables that do not necessitate opening, include information on the types of objects placed on their surfaces.
Open Success	Provide feedback confirming the successful opening to the target object, along with detailed information regarding the types and positions of items contained within it.
Move Success	Provide feedback indicating successful displacement by the specified distances along both the x-axis and y-axis.
Pick Success	Return feedback that the target object has been successfully picked.
Place Success	Provide feedback on the successful placement of the target object at the specified location.
Feedback on failed actions	
Navigation Failed	(1) The starting or ending point of global path planning is deemed invalid if it falls on an obstacle or exceeds the map boundary; (2) The target object for navigation is considered invalid if it does not exist in the scene graph or does not support navigation; (3) A discrepancy greater than an acceptable threshold between the robot's current pose and the target pose can lead to failure.
Open Failed	(1) The target object is either already in an open state or cannot be opened; (2) The target object is positioned beyond the operational range of the robot.
Move Failed	The feedback type aligns with "navigation failed" due to the invocation of the API responsible for the <i>navigate</i> command.
Pick Failed	(1) The gripper is grabbing with other objects; (2) The scene graph lacks information about the object to be grasped; (3) The initial or target state is invalid during the verification of the robot arm planning algorithm; (4) The distance between the robot arm's end effector and the target object exceeds the allowable threshold. Return this distance, and for mobile manipulation robots, also provide the relative distance between the base and the target object along the x and y axes.
Place Failed	(1) The gripper is empty; (2) The object to be placed does not match the object currently being grasped; (3) The object to be placed has not been placed at the target location.
Common feedback	
Target Task Status	Only available for manipulation robots. (1) In the "pack objects" task, the feedback refers to the types of objects present in the tray; (2) In the "sort solids" task, it provides information regarding the shapes and colors of solids placed on panels of different colors; (3) In the "make sandwich" task, the feedback describes the types of ingredients arranged on the cutting board from bottom to top.

- *Pack Objects*: The objective of this task is to evaluate the robot's fundamental picking and placing capabilities. The robot is provided with a list of objects and must accurately place each object into a designated tray. The objects include *apple, fork, soap, toy duck, phone, bottle, book*, etc.
- *Sort solids*: In addition to evaluating the robot's basic pick-and-place capabilities, this task also requires the robot to perform color matching. The robot must accurately place solids, each of a different color, onto the corresponding colored panels on the table. The colors include *red, blue, pink, green, yellow, and purple*.
- *Make Sandwich*: This task further tests the robot's ability to stack objects in a specific order. The robot is tasked with assembling sandwiches of varying flavors based on a given menu, requiring it to sequentially stack the sandwich ingredients. The ingredients include *bread slices, ham, bacon, tomato, cucumber, cheese, and beef patties*.

Additionally, in the three types of tasks outlined above, the mobile robot (David) is responsible for the efficient exploration of the environment. The mobile manipulation robot (Alice), beyond exploration, is tasked with opening furniture and transporting objects. The manipulation robot (Bob) handles operations on the tabletop. Alice is not capable

of directly executing the final manipulation tasks; all tabletop operations are performed by Bob.

We introduce four metrics to evaluate the performance of different methods:

- *Success rate (Succ)*: The rate of episodes in which the robot completes the full task, meaning all objects are placed in the correct positions.
- *Partial success rate (PS)*: The average ratio of correctly placed objects to the total number of objects in each episode. The rationale for designing this metric is that, within a task episode, an increase in the number of objects and the length of the sequence generally makes the task more challenging. However, with more objects, the impact of a single failure tends to be smaller. Conversely, in tasks involving fewer objects, the consequences of a single failure are more significant.
- *Average temporal steps (TS)*: the average number of temporal steps in all episodes.
- *Average action steps (AS)*: the average number of action steps in all episodes, excluding *wait* actions. The robot's capability to execute *wait* actions at optimal moments can significantly reduce energy consumption, demonstrating that the policy has a comprehensive understanding of the entire task.

## B. Experiments on simulation scenarios

TABLE II  
EVALUATION RESULTS ON OUR ENVIRONMENT AND TASKS

Methods	Succ(%) $\uparrow$	PS(%) $\uparrow$	TS $\downarrow$	AS $\downarrow$
Make Sandwich				
MHRC w. GPT-3.5-turbo	0.00	0.14	50.00	41.22
MHRC w. Llama-3.1-8B	0.00	0.12	50.00	46.00
MHRC w. GPT-4o	<b>0.75</b>	<b>0.81</b>	<b>38.89</b>	<b>31.11</b>
Sort Solids				
MHRC w. GPT-3.5-turbo	0.00	0.17	50.00	39.89
MHRC w. Llama-3.1-8B	0.00	0.14	50.00	44.33
MHRC w. GPT-4o	<b>0.83</b>	<b>0.88</b>	<b>34.56</b>	<b>27.33</b>
Pack Objects				
MHRC w. GPT-3.5-turbo	0.00	0.25	50.00	37.67
MHRC w. Llama-3.1-8B	0.00	0.20	50.00	40.33
MHRC w. GPT-4o	<b>0.92</b>	<b>0.96</b>	<b>32.67</b>	<b>25.00</b>

1) *Comparative experiments with different LLMs:* In our framework, we evaluated various LLMs, including the closed-source models GPT-3.5-turbo and GPT-4o, as well as the open-source Llama-3.1-8B. Tab. II summarizes their performance across multiple metrics. Each task was tested in three distinct scenarios, with four trials per scenario. The number of objects per trial was set to 3, 4, 5, and 6, respectively, creating task sequences of varying lengths. We ensured that each task involved objects requiring exploration and transportation by the robots, as well as objects positioned on a table but beyond the reach of the tabletop robots. The maximum step count for each experiment was limited to 50, and all models were tested with a temperature setting of 0.5. Experimental results showed that, aside from GPT-4o, the other models struggled to complete full tasks. They were only able to successfully grasp objects near the robotic arm and frequently made logically inconsistent decisions, such as issuing an open command far from the refrigerator.

2) *Ablation study:* Our multi-heterogeneous robot collaboration framework comprises several key modules, with Tab. III detailing the performance outcomes when various components are ablated. The experimental results highlight the necessity of both feedback and history, as they play a pivotal role in determining task execution success. Without access to history information, the model may lose track of prior actions, leading to behaviors such as re-exploring the same location or forgetting previously received requests. This results in unnecessary steps due to redundant communication. Additionally, the manipulation robot often focuses only on nearby objects, failing to collaborate effectively on objects that require coordination.

The absence of feedback proves even more detrimental, causing a significant decrease in task success rates and the emergence of various errors. For example, the mobile manipulation robot might not detect a failed grasp attempt and proceed with subsequent actions regardless, even sending incorrect status updates to the manipulation robot, claiming successful placement of objects on the table. Similarly, the manipulation robot may limit its actions to nearby objects, ignoring the broader task requirements. Furthermore, the

TABLE III  
EVALUATION RESULTS ON OUR ENVIRONMENT AND TASKS

Methods	Succ(%) $\uparrow$	PS(%) $\uparrow$	TS $\downarrow$	AS $\downarrow$
Make Sandwich				
MHRC w.o. feedback	0.00	0.08	50.00	48.56
MHRC w.o. history	0.00	0.48	50.00	46.89
MHRC w.o. mobile robot	0.69	0.78	44.67	36.33
MHRC w. GPT-4o	<b>0.75</b>	<b>0.81</b>	<b>38.89</b>	<b>31.11</b>
Sort Solids				
MHRC w. feedback	0.00	0.10	50.00	46.22
MHRC w. history	0.00	0.56	50.00	47.44
MHRC w.o. mobile robot	0.81	0.84	40.50	31.17
MHRC w. GPT-4o	<b>0.83</b>	<b>0.88</b>	<b>34.56</b>	<b>27.33</b>
Pack Objects				
MHRC w. feedback	0.00	0.14	50.00	46.00
MHRC w. history	0.00	0.61	50.00	48.33
MHRC w.o. mobile robot	<b>0.92</b>	0.94	37.00	28.33
MHRC w. GPT-4o	<b>0.92</b>	<b>0.96</b>	<b>32.67</b>	<b>25.00</b>

robots operating within the framework without feedback and history exhibited significantly higher AS values, suggesting that these systems lacked a comprehensive understanding of the overall task, resulting in excessive redundant actions.

While theoretically, Alice and Bob alone are sufficient to complete all the experimental tasks, we explored the impact of removing David from the framework to assess the resultant performance. As demonstrated by the experimental results in Tab. III, David plays a crucial role in enhancing task efficiency.

## V. CONCLUSIONS AND FUTURE WORK

In this paper, we present MHRC, a novel framework that leverages LLMs for decentralized collaboration among heterogeneous multi-robot systems. We designed tailored feedback mechanisms for different robot types, enabling them to replan based on real-time feedback, thus improving task success rates. Our approach demonstrated effective decision-making and planning, as validated through a series of simulated tasks.

In future work, two aspects can be further explored: (1) The integration of additional heterogeneous robots to address more complex environments and tasks, such as deploying legged robots to navigate uneven terrains like stairs, and drones to explore aerial targets. (2) Although the current framework can complete tasks within a limited time, it remains inefficient, with redundant actions. Thus, developing a more optimized collaboration framework, along with enhanced CoT approaches, will facilitate more efficient robot cooperation.

## APPENDIX

### A. Detailed prompt templates

Due to the extensive content of the prompt template and the constraints of page space, please visit our website for detailed information. [https://github.com/WenhaoYu1998/ICRA25/blob/main/ICRA25\\_MHRC\\_appendix.pdf](https://github.com/WenhaoYu1998/ICRA25/blob/main/ICRA25_MHRC_appendix.pdf)

## REFERENCES

- [1] Y. Rizk, M. Awad, and E. W. Tunstel, "Cooperative heterogeneous multi-robot systems: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 2, pp. 1–31, 2019.
- [2] O. Shorinwa, T. Halsted, J. Yu, and M. Schwager, "Distributed optimization methods for multi-robot systems: Part 2—a survey," *IEEE Robotics & Automation Magazine*, 2024.
- [3] A. Kolling and S. Carpin, "Multirobot cooperation for surveillance of multiple moving targets-a new behavioral approach," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006*. IEEE, 2006, pp. 1311–1316.
- [4] X. Liu, P. Li, W. Yang, D. Guo, and H. Liu, "Leveraging large language model for heterogeneous ad hoc teamwork collaboration," *arXiv preprint arXiv:2406.12224*, 2024.
- [5] Z. Mandi, S. Jain, and S. Song, "Roco: Dialectic multi-robot collaboration with large language models," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 286–299.
- [6] C. Amato, G. Chowdhary, A. Geramifard, N. K. Üre, and M. J. Kochenderfer, "Decentralized control of partially observable markov decision processes," in *52nd IEEE Conference on Decision and Control*. IEEE, 2013, pp. 2398–2405.
- [7] D. S. Bernstein, S. Zilberman, and N. Immerman, "The complexity of decentralized control of markov decision processes," 2013. [Online]. Available: <https://arxiv.org/abs/1301.3836>
- [8] J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. Le, and D. Zhou, "Chain-of-thought prompting elicits reasoning in large language models," 2023. [Online]. Available: <https://arxiv.org/abs/2201.11903>
- [9] K. Greff, F. Belletti, L. Beyer, C. Doersch, Y. Du, D. Duckworth, D. J. Fleet, D. Gnanapragasam, F. Golemo, C. Herrmann, et al., "Kubric: A scalable dataset generator," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 3749–3761.
- [10] P. A. Jansen, "Visually-grounded planning without vision: Language models infer detailed plans from high-level instructions," 2020. [Online]. Available: <https://arxiv.org/abs/2009.14259>
- [11] P. Sharma, A. Torralba, and J. Andreas, "Skill induction and planning with latent language," 2022. [Online]. Available: <https://arxiv.org/abs/2110.01517>
- [12] W. Huang, P. Abbeel, D. Pathak, and I. Mordatch, "Language models as zero-shot planners: Extracting actionable knowledge for embodied agents," in *International conference on machine learning*. PMLR, 2022, pp. 9118–9147.
- [13] S. Li, X. Puig, C. Paxton, Y. Du, C. Wang, L. Fan, T. Chen, D.-A. Huang, E. Akyürek, A. Anandkumar, et al., "Pre-trained language models for interactive decision-making," *Advances in Neural Information Processing Systems*, vol. 35, pp. 31 199–31 212, 2022.
- [14] M. Ahn, A. Brohan, N. Brown, Y. Chebotar, O. Cortes, B. David, C. Finn, C. Fu, K. Gopalakrishnan, K. Hausman, et al., "Do as i can, not as i say: Grounding language in robotic affordances," *arXiv preprint arXiv:2204.01691*, 2022.
- [15] A. Zeng, M. Attarian, B. Ichter, K. Choromanski, A. Wong, S. Welker, F. Tombari, A. Purohit, M. Ryoo, V. Sindhwani, J. Lee, V. Vanhoucke, and P. Florence, "Socratic models: Composing zero-shot multimodal reasoning with language," 2022. [Online]. Available: <https://arxiv.org/abs/2204.00598>
- [16] I. Singh, V. Blukis, A. Mousavian, A. Goyal, D. Xu, J. Tremblay, D. Fox, J. Thomason, and A. Garg, "Progrompt: Generating situated robot task plans using large language models," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 11 523–11 530.
- [17] J. Liang, W. Huang, F. Xia, P. Xu, K. Hausman, B. Ichter, P. Florence, and A. Zeng, "Code as policies: Language model programs for embodied control," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 9493–9500.
- [18] W. Huang, C. Wang, R. Zhang, Y. Li, J. Wu, and L. Fei-Fei, "Voxposer: Composable 3d value maps for robotic manipulation with language models," 2023. [Online]. Available: <https://arxiv.org/abs/2307.05973>
- [19] W. Huang, F. Xia, T. Xiao, H. Chan, J. Liang, P. Florence, A. Zeng, J. Tompson, I. Mordatch, Y. Chebotar, P. Sermanet, N. Brown, T. Jackson, L. Luu, S. Levine, K. Hausman, and B. Ichter, "Inner monologue: Embodied reasoning through planning with language models," 2022. [Online]. Available: <https://arxiv.org/abs/2207.05608>
- [20] T. B. Brown, "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.
- [21] M. Chen, J. Tworek, H. Jun, Q. Yuan, H. P. D. O. Pinto, J. Kaplan, H. Edwards, Y. Burda, N. Joseph, G. Brockman, et al., "Evaluating large language models trained on code," *arXiv preprint arXiv:2107.03374*, 2021.
- [22] Y. Liu, "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019.
- [23] N. Reimers, "Sentence-bert: Sentence embeddings using siamese bert-networks," *arXiv preprint arXiv:1908.10084*, 2019.
- [24] P. Stone and M. Veloso, "Multiagent systems: A survey from a machine learning perspective," *Autonomous Robots*, vol. 8, pp. 345–383, 2000.
- [25] Y. Koga and J.-C. Latombe, "On multi-arm manipulation planning," in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, 1994, pp. 945–952 vol.2.
- [26] A. Dorri, S. S. Kanhere, and R. Jurdak, "Multi-agent systems: A survey," *Ieee Access*, vol. 6, pp. 28 573–28 593, 2018.
- [27] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," 2011. [Online]. Available: <https://arxiv.org/abs/1105.1186>
- [28] A. Dobson and K. E. Bekris, "Planning representations and algorithms for prehensile multi-arm manipulation," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 6381–6386.
- [29] J. Li, H. Modares, T. Chai, F. L. Lewis, and L. Xie, "Off-policy reinforcement learning for synchronization in multiagent graphical games," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2434–2445, 2017.
- [30] H. Zhang, J. H. Park, D. Yue, and X. Xie, "Finite-horizon optimal consensus control for unknown multiagent state-delay systems," *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 402–413, 2020.
- [31] H. Zhang, W. Du, J. Shan, Q. Zhou, Y. Du, J. B. Tenenbaum, T. Shu, and C. Gan, "Building cooperative embodied agents modularly with large language models," *arXiv preprint arXiv:2307.02485*, 2023.
- [32] X. Guo, K. Huang, J. Liu, W. Fan, N. Vélez, Q. Wu, H. Wang, T. L. Griffiths, and M. Wang, "Embodied llm agents learn to cooperate in organized teams," *arXiv preprint arXiv:2403.12482*, 2024.
- [33] Y. Q. Chen and Z. Wang, "Formation control: a review and a new consideration," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, pp. 3181–3186.
- [34] N. Lissandrini, C. K. Verginis, P. Roque, A. Cenedese, and D. V. Dimarogonas, "Decentralized nonlinear mpc for robust cooperative manipulation by heterogeneous aerial-ground robots," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 1531–1536.
- [35] Y. Yue, M. Wen, Y. Putra, M. Wang, and D. Wang, "Tightly-coupled perception and navigation of heterogeneous land-air robots in complex scenarios," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 10 052–10 058.
- [36] D. W. Haldane, P. Fankhauser, R. Siegwart, and R. S. Fearing, "Detection of slippery terrain with a heterogeneous team of legged robots," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 4576–4581.
- [37] D. Wallace, Y. He, J. Chagas Vaz, L. Georgescu, and P. Y. Oh, "Multimodal teleoperation of heterogeneous robots within a construction environment," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 2698–2705.
- [38] Y. Zhao, B. Huang, J. Yu, and Q. Zhu, "Stackelberg strategic guidance for heterogeneous robots collaboration," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, May 2022, p. 4922–4928. [Online]. Available: <http://dx.doi.org/10.1109/ICRA46639.2022.9811678>
- [39] M. Hu, Y. Mu, X. Yu, M. Ding, S. Wu, W. Shao, Q. Chen, B. Wang, Y. Qiao, and P. Luo, "Tree-planner: Efficient close-loop task planning with large language models," 2024. [Online]. Available: <https://arxiv.org/abs/2310.08582>
- [40] W. Huang, F. Xia, T. Xiao, H. Chan, J. Liang, P. Florence, A. Zeng, J. Tompson, I. Mordatch, Y. Chebotar, P. Sermanet, N. Brown, T. Jackson, L. Luu, S. Levine, K. Hausman, and B. Ichter, "Inner monologue: Embodied reasoning through planning with language models," in *arXiv preprint arXiv:2207.05608*, 2022.
- [41] C. Ming, J. Lin, P. Fong, H. Wang, X. Duan, and J. He, "Hicrisp: A hierarchical closed-loop robotic intelligent self-correction planner," *arXiv preprint arXiv:2309.12089*, 2023.

- [42] P. Zhi, Z. Zhang, M. Han, Z. Zhang, Z. Li, Z. Jiao, B. Jia, and S. Huang, “Closed-loop open-vocabulary mobile manipulation with gpt-4v,” *arXiv preprint arXiv:2404.10220*, 2024.
- [43] A. Szot, K. Yadav, A. Clegg, V.-P. Berges, A. Gokaslan, A. Chang, M. Savva, Z. Kira, and D. Batra, “Habitat rearrangement challenge 2022,” 2022. [Online]. Available: [https://aihabitat.org/challenge/rearrange\\_2022](https://aihabitat.org/challenge/rearrange_2022)
- [44] J. Gu, D. S. Chaplot, H. Su, and J. Malik, “Multi-skill mobile manipulation for object rearrangement,” *arXiv preprint arXiv:2209.02778*, 2022.
- [45] A. Makhal and A. K. Goins, “Reuleaux: Robot base placement by reachability analysis,” in *2018 second IEEE international conference on robotic computing (IRC)*. IEEE, 2018, pp. 137–142.
- [46] B. Shao, Y. Ding, X. Wang, X. Xie, F. Gu, J. Luo, and C. Chen, “Moma-pos: Where should mobile manipulators stand in cluttered environment before task execution?” *arXiv preprint arXiv:2403.19940*, 2024.
- [47] Y. Ding, X. Zhang, C. Paxton, and S. Zhang, “Task and motion planning with large language models for object rearrangement,” in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 2086–2092.
- [48] Y. Ding, X. Zhang, S. Amiri, N. Cao, H. Yang, A. Kaminski, C. Esselink, and S. Zhang, “Integrating action knowledge and llms for task planning and situation handling in open worlds,” *Autonomous Robots*, vol. 47, no. 8, pp. 981–997, 2023.
- [49] K. Kurzer, “Path planning in unstructured environments: A real-time hybrid a\* implementation for fast and deterministic path generation for the kth research concept vehicle,” Master’s thesis, 2016.
- [50] D. Fox, W. Burgard, and S. Thrun, “The dynamic window approach to collision avoidance,” *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, pp. 23–33, 1997.
- [51] W. Yu, J. Peng, Q. Qiu, H. Wang, L. Zhang, and J. Ji, “Pathrl: An end-to-end path generation method for collision avoidance via deep reinforcement learning,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 9278–9284.
- [52] W. Yu, J. Peng, H. Yang, J. Zhang, Y. Duan, J. Ji, and Y. Zhang, “Ldp: A local diffusion planner for efficient robot navigation and collision avoidance,” *arXiv preprint arXiv:2407.01950*, 2024.
- [53] J. D. Gammell, S. S. Srinivasa, and T. D. Barfoot, “Batch informed trees (bit\*): Sampling-based optimal planning via the heuristically guided search of implicit random geometric graphs,” in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 3067–3074.