

Phase-Adaptive Large Language Models for Multi-Robot Task Allocation in Dynamic Construction Environments

Shyam Prasad Reddy Kaitha¹ and Hongrui Yu^{2*}

¹PhD. Student, Dept. of Civil and Environmental Engineering, Virginia Tech, United States of America

²Assistant Professor, Dept. of Civil and Environmental Engineering, Virginia Tech, United States of America

skaitha@vt.edu, hryu42@vt.edu

Abstract -

Construction robotics requires effective coordination strategies for heterogeneous robot teams performing long-horizon and complex tasks. Traditional approaches using Dynamic Programming and Reinforcement Learning (RL) have dominated multi-robot task allocation, yet emerging Large Language Model (LLM) solutions demonstrate a more productive allocation scheme. To further improve LLM-based multi-robot task allocation for construction, the authors present a phase-adaptive strategy for LLM-driven task allocation that addresses the core challenge between maximizing successful task execution and maintaining balanced workload distribution. Our approach adjusts allocation priorities dynamically: early project stages emphasize task completion to establish operational foundations, while later stages prioritize equitable work distribution as adjustment opportunities decrease. The authors have also provided systematic evaluation against RL-based methods using a benchmark dataset. Our framework achieves a higher and more robust performance comparable to RL methods while delivering workload balance. Results demonstrate that LLM-based coordination with adaptive phase goal prioritization offers a better alternative to RL, providing interpretability and flexibility without intensive model retraining efforts for construction automation applications.

Keywords -

Construction Robotics; Multi-Robot Task Allocation; Large Language Models; Phase-Adaptive Allocation

1 Introduction

The construction industry faces critical skilled labor shortages that significantly impact project performance. Recent analysis reveals that the industry requires an estimated 501,000 additional workers beyond normal hiring to meet demand, a deficit that leads to substantial cost overruns, schedule delays, and compromised safety outcomes [1] [2]. These workforce constraints necessitate automation solutions that can augment human capabilities while maintaining productivity and safety standards. Construction robotics presents a viable approach to ad-

dress these challenges, particularly through coordinated multi-robot systems capable of executing complex, long-horizon tasks that require diverse specialized capabilities such as material handling, precision assembly, and quality inspection [3].

Effective deployment of construction robots requires sophisticated coordination frameworks that can allocate tasks across heterogeneous robot teams. Multi-Robot Task Allocation (MRTA) addresses this challenge by decomposing complex workflows into manageable subtasks and assigning them to appropriate robots based on capability matching and environmental constraints [4]. The coordination challenge extends beyond static task distribution: construction sites are inherently dynamic environments where task requirements evolve, spatial constraints shift, and the performance-fairness trade-off changes across project phases [5][6].

Construction environments present particularly demanding MRTA scenarios characterized by multi-robot, multi-task, and time-extended allocation requirements with dynamic spatial constraints and evolving temporal dependencies throughout project execution. Traditional approaches to MRTA rely primarily on mathematical optimization methods and reinforcement learning (RL) techniques. While these methods provide rigorous performance guarantees, they face significant limitations in construction contexts. RL-based coordination requires extensive trial-and-error training that poses safety risks on active construction sites where coordination failures could endanger workers or damage equipment [7].

Additionally, optimization algorithms often require complete problem specification at planning time, limiting their adaptability to the dynamic conditions characteristic of construction environments.

Recent advances in Large Language Models (LLMs) offer an alternative paradigm for multi-robot coordination through natural language reasoning and zero-shot task planning capabilities. LLM-based frameworks have demonstrated potential for robot task allocation through structured prompting approaches that eliminate the need for extensive training data or model fine-tuning [8]. However, a critical gap exists: no prior research has systemati-

cally benchmarked LLM-driven task allocation against established optimization methods for construction robot coordination, despite growing evidence that LLMs can serve as effective multi-robot planners [9]. Furthermore, existing frameworks apply static allocation priorities throughout task sequences, failing to address the temporal evolution of the performance-fairness trade-off as projects progress and re-balancing opportunities diminish.

This research addresses these gaps by introducing the LangGraph-based Task Allocation Agent (LTAA) framework. The framework employs a phase-adaptive allocation strategy that dynamically shifts priorities from execution feasibility in early project stages (when flexibility is high) to workload equity in later stages (when rebalancing opportunities decrease). The specific objectives are to: (1) develop and validate this phase-adaptive strategy for construction robot coordination, and (2) conduct systematic performance benchmarking of the LLM-based approach against traditional reinforcement learning methods (Q-learning, Deep Q-Networks) using construction-related tasks from the TEACH dataset.

The remainder of this paper is organized as follows: Section 2 reviews related work, Section 3 presents the LTAA framework methodology, Section 4 describes the experimental setup, Section 5 presents results, and Section 6 concludes with findings and future directions.

2 Related works

2.1 Multi-Robot Task Allocation for Construction

Multi-Robot Task Allocation (MRTA) provides the theoretical foundation for coordinating robotic teams through systematic task distribution considering resource constraints, temporal requirements, and robot capability heterogeneity. Gerkey and Mataric [4] established the foundational MRTA taxonomy characterizing problems along three dimensions: single-task versus multi-task robots, single-robot versus multi-robot tasks, and instantaneous versus time-extended assignments. Construction coordination falls into the most complex category multi-robot, multi-task, time-extended allocation (MT-MR-TA) requiring sophisticated reasoning to handle heterogeneous capabilities, diverse task requirements, and temporal dependencies across project phases.

Multi-robot task allocation has gained increasing attention in construction automation as projects require coordination of heterogeneous robot teams with specialized capabilities. Ye et al. [10] developed a two-stage genetic algorithm approach for construction robot task allocation, demonstrating improved performance over traditional single-stage optimization methods. Chakraborty et al. [11] provided a comprehensive review of optimization techniques for multi-robot task allocation, identifying key

challenges in handling dynamic environments and heterogeneous robot capabilities characteristic of construction sites. These studies establish that construction MRTA requires sophisticated approaches capable of managing complex constraints including spatial dependencies, temporal sequencing, and capability matching across diverse robot teams.

2.2 Traditional Optimization Approaches

Recent advances in learning-based methods have addressed some limitations of classical optimization techniques for construction robot coordination. Lee et al. [7] introduced a digital twin-driven deep reinforcement learning framework for adaptive task allocation in robotic construction, demonstrating that DRL can handle dynamic construction environments through continuous learning. However, their approach requires extensive training data and faces challenges in sim-to-real transfer for safety-critical construction applications. Chen et al. [12] developed a multi-objective ant colony optimization algorithm for heterogeneous unmanned aerial vehicle task allocation, showing effectiveness in handling multiple competing objectives. While these learning-based methods offer adaptability, they share common limitations: extensive computational requirements for training, limited interpretability of decision-making processes, and difficulty incorporating domain-specific construction constraints without significant algorithm modification.

2.3 LLM-based Coordination Approaches

Large language models have emerged as an alternative paradigm for robot coordination through natural language reasoning capabilities that eliminate extensive training requirements. Kannan et al. [8] developed the SMART-LLM framework for LLM-driven multi-robot task planning, demonstrating that language models could effectively perform task decomposition and allocation using structured prompts, achieving 70% success rates in household robotics scenarios while maintaining interpretability through natural language reasoning. However, SMART-LLM was evaluated primarily in static household environments and did not address the dynamic replanning requirements characteristic of construction projects. The authors' engagement with this framework further revealed that static allocation priorities struggle with the performance-fairness trade-off as construction projects progress a limitation that directly motivated the phase-adaptive design of LTAA.

Deng et al. [13] developed an integrated framework combining digital twins, optimization backends, and LLM-driven interpretation for dynamic construction task rescheduling, achieving high accuracy in constraint extraction but focusing primarily on reactive rescheduling rather

than initial allocation optimization. Prieto and García de Soto [14] introduced collaborative multi-agent LLM architectures employing Planner and Supervisor agents for construction task allocation, demonstrating improved reliability through agent coordination. Kim et al. [15] developed frameworks integrating BIM-based knowledge with natural language dialogue for construction applications, showing promise for human-robot interaction. However, these LLM-based approaches lack systematic performance benchmarking against established optimization methods and do not address temporal evolution of allocation priorities across project phases, representing critical gaps for construction deployment where both performance validation and phase-specific adaptation are essential.

3 METHODOLOGY

3.1 LTAA Framework Overview

The LangGraph-based Task Allocation Agent (LTAA) framework addresses construction robot task allocation through an LLM-driven coordination approach that accepts pre-decomposed construction tasks and generates capability-aware robot assignments. The framework operates through a sequential workflow orchestrated by LangGraph, consisting of three primary components: input state management, phase-adaptive allocation logic, and allocation validation. Figure 1 illustrates the complete framework architecture. The system receives as input a construction task with specified feature requirements (e.g., heavy lifting, precision assembly, careful handling), the current state of robot workloads, and project progression metrics. The allocation agent then evaluates robot-task compatibility through probabilistic success modeling while simultaneously assessing workload distribution to determine appropriate allocation priorities based on the current project phase. Dynamic prompting enables the framework to adjust its reasoning strategy without requiring model fine-tuning, constructing phase-specific prompts that incorporate current workload context and priority guidance. Upon generating an allocation decision, the system validates output format and constraint satisfaction before finalizing the assignment and updating workload tracking for subsequent allocations.

3.2 Phase-Adaptive Allocation Strategy

Construction projects evolve temporally, requiring different allocation priorities at different stages. Early-phase allocation must emphasize execution feasibility to establish operational baselines, while late-phase allocation should prioritize workload equity as opportunities for rebalancing diminish[5][6]. The phase-adaptive strategy addresses this temporal evolution through systematic priority adjustment across three project phases.

Phase Detection Mechanism: The framework classifies allocation cycles into three phases based on task completion progress: Early (0-33%), Middle (34-66%), and Late (67-100%). Algorithm 1 presents the phase detection logic. For each allocation cycle, the system computes the progress ratio as the number of allocated tasks divided by total tasks, then maps this ratio to the appropriate phase. The early phase prioritizes success rate maximization when flexibility is highest. The middle phase begins balancing success and equity as workload disparities emerge. The late phase emphasizes equitable completion as remaining opportunities for redistribution decrease.

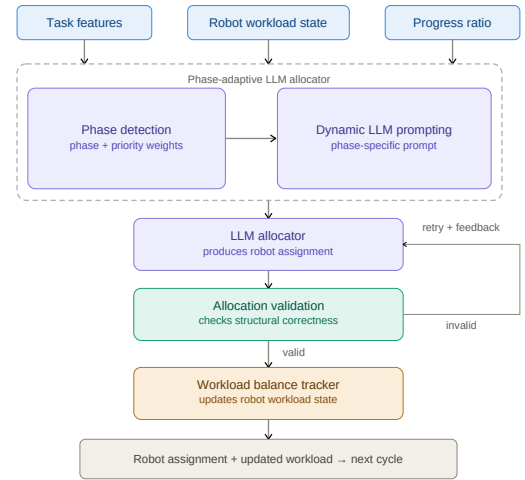


Figure 1. LangGraph based Task Allocation Agent Framework Architecture

Algorithm 1 Phase Detection for Task Allocation

Require: Task sequence T , current allocation index i

Ensure: Allocation phase, balance urgency

```

1: Compute number of allocated tasks  $c_i$ 
2: Compute total number of tasks  $C$ 
3: if  $C = 0$  then
4:    $\text{phase} \leftarrow \text{Early}$ 
5:    $\text{balance\_urgency} \leftarrow 0$ 
6: else
7:   Compute progress ratio  $p \leftarrow \frac{c_i}{C}$ 
8:   if  $p < 0.33$  then
9:      $\text{phase} \leftarrow \text{Early}$ 
10:   else if  $0.33 \leq p < 0.67$  then
11:      $\text{phase} \leftarrow \text{Middle}$ 
12:   else
13:      $\text{phase} \leftarrow \text{Late}$ 
14:   end if
15:    $\text{balance\_urgency} \leftarrow p$ 
16: end if
17: return phase, balance_urgency

```

Priority Shifting Logic: The framework implements phase-specific decision frameworks that guide allocation

reasoning. In the early phase, the system accepts initial workload imbalance in favor of establishing high success rates, reasoning that subsequent allocations provide opportunities for correction. The prompt instructs the LLM to “prioritize success rate more heavily” and “accept some imbalance initially.” In the middle phase, the framework introduces balanced consideration of both objectives, instructing the system to “weigh trade-offs” and “consider criticality, imbalance severity, phase, and future tasks.” The late phase elevates workload equity, guiding the LLM to “favor the underutilized robot” when success gaps are moderate (within 10-20%), while maintaining performance for critical allocations. This graduated approach enables the framework to achieve competitive task completion rates while maintaining superior workload distribution compared to static optimization strategies.

3.3 Workload Balance Quantification

Effective workload balancing requires quantitative metrics for measuring distribution fairness and detecting imbalance severity[16]. The framework tracks the number of tasks assigned to each robot and computes deviation from an ideal uniform workload distribution. For a team of n robots with a total of T assigned tasks, the target workload per robot is defined as:

$$W_{\text{target}} = \frac{T}{n}$$

The workload deviation for robot r is computed as:

$$\delta_r = W_r - W_{\text{target}}$$

where W_r denotes the number of tasks assigned to robot r .

Overall workload fairness is quantified using a balance score B defined on a 0–100 scale:

$$B = 100 \times \left(1 - \frac{\max_r |\delta_r|}{W_{\text{target}}} \right) \quad (1)$$

Here, $\max_r |\delta_r|$ represents the maximum absolute workload deviation among all robots, capturing the worst-case imbalance in the system. A balance score of 100 indicates perfect workload equity, while lower values correspond to increasing imbalance.

To support adaptive decision-making, imbalance severity is classified into four levels based on the ratio $\max_r |\delta_r|/W_{\text{target}}$: low (< 0.2), moderate ($0.2 \leq \text{ratio} < 0.4$), high ($0.4 \leq \text{ratio} < 0.6$), and critical (≥ 0.6). These severity levels inform phase-specific allocation strategies.

3.4 Probabilistic Robot Capability Modeling

Construction tasks require diverse capabilities, and robot teams typically comprise heterogeneous units with

specialized strengths. The framework models three robot types representing common construction robotics profiles: Light robots excel at precision and careful manipulation (optimal for assembly, inspection, delicate handling), Medium robots provide balanced generalist capabilities (suitable for mixed operations and general-purpose tasks), and Heavy robots specialize in force application (optimal for heavy lifting, material handling, construction operations). Table 1 summarizes robot specifications and optimal applications.

Table 1. Robot Specifications and Optimal Applications

Robot	Specialization	Optimal Applications
Light	Precision Specialist	Assembly, inspection, delicate manipulation
Medium	Balanced Generalist	Mixed operations, general-purpose tasks
Heavy	Force Specialist	Heavy lifting, material handling, construction

Each construction task is characterized by required features representing operational demands[17]: Heavy tasks require substantial force or endurance (e.g., material transport, positioning components), Dexterous tasks demand high precision and fine motor skills (e.g., intricate detailing, sealing joints), and Careful tasks emphasize cautious handling to avoid damage (e.g., handling fragile materials, aligning precision equipment). Tasks may require multiple features simultaneously, reflecting the multi-faceted nature of construction operations.

For each task t defined by a set of required features $F_t = \{f_1, f_2, \dots, f_n\}$, the framework computes an aggregate success rate $S(r, t)$ for each robot r as follows:

$$S(r, t) = \frac{1}{|F_t|} \sum_{f \in F_t} M[f][r] \quad (2)$$

where $M[f][r]$ denotes the success probability of robot r for feature f , and $|F_t|$ represents the number of required features. This aggregation provides a single success metric that the allocation agent uses in combination with workload balance considerations to make informed assignment decisions. The probabilistic modeling enables the framework to reason about task-robot compatibility without requiring explicit optimization formulations, supporting the LLM’s natural language reasoning capabilities while maintaining quantitative performance assessment.

4 Experimental Setup

4.1 Task Dataset

The proposed phase-adaptive task allocation approach is evaluated using long-horizon task sequences derived

from the TEACH dataset[18] in a computational simulation environment; physical robot deployment is outside the scope of this study, which focuses on validating allocation decision-making logic. TEACH provides task instructions that require sequential decision-making and heterogeneous operational capabilities, making it suitable for studying multi-robot task allocation under construction-like conditions. This domain-abstraction approach frames allocation as a capability–task feature matching problem, consistent with the capability decomposition principle proposed by Ranz et al.[19] where tasks are decomposed into discrete capability requirements and matched to agents based on their capability profiles. The structural properties of TEACH dataset specifically its sequential long-horizon task organization, multi-step decision dependencies, and heterogeneous operational demands across tasks align with the coordination challenges characteristic of construction workflows, positioning both within the MT-MR-TA category of the Gerkey-Matarić taxonomy [4], making TEACH suitable for evaluating the allocation logic independently of the physical domain. Consistent with simulation-based benchmarking in recent heterogeneous multi-robot research [20]. Tasks are abstracted using feature-based requirements reflecting common construction operations, Heavy (material transport, component positioning), Dexterous (fastener insertion, joint sealing), and Careful (fragile material handling, sensor alignment) to reflect operational demands characteristic of construction sites.

4.2 Robot Team Configuration

The experimental setup considers a heterogeneous team of three representative construction robot types: light, medium, and heavy robots. This three-robot heterogeneous configuration follows established experimental conventions in multi-robot task allocation research, where capability diversity is prioritized over team size to isolate algorithmic behavior[4][21]. Each robot operates independently under a single-task assignment model (ST-SR in the Gerkey-Matarić taxonomy), receiving sequential task assignments without inter-robot collaboration on individual tasks. Light robots are optimized for precision and careful manipulation, medium robots provide balanced general-purpose capabilities, and heavy robots specialize in force-intensive operations. Each robot type is associated with distinct success probabilities across task features, capturing capability specialization commonly observed in construction robotics. Specifically, these robot classes correspond to typical construction platforms: Light robots represent collaborative arms (e.g., UR5e, Franka Emika Panda; payload 5–15 kg), Medium robots represent versatile mobile manipulators (payload 15–50 kg), and Heavy robots represent force-intensive construction equipment

(e.g., Hadrian X-class; payload 50–200+ kg).

Table 2. Robot Capability Success Probabilities Across Task Features

Robot	Careful	Dexterous	Heavy
Light	0.90	0.80	0.50
Medium	0.70	0.60	0.70
Heavy	0.50	0.40	0.90

Robot capability differences are represented using probabilistic success rates across task features, as summarized in Table 2. These values denote $P(\text{success} / \text{robot type } r, \text{task feature } f)$, derived from canonical capability profiles in multi-robot task allocation literature [17][19][22] rather than direct hardware measurement. Light robots receive high probabilities for precision-dependent features consistent with collaborative arm characteristics; heavy robots for force-intensive operations; and medium robots reflect balanced generalist profiles. The success probabilities are used to compute task–robot compatibility via Equation (2) and remain fixed across all allocation methods to ensure fair comparison.

4.3 Baseline Methods

The proposed approach is benchmarked against two reinforcement learning-based task allocation methods: Q-learning and Deep Q-Networks (DQN). These methods represent established learning-based baselines capable of handling stochastic task outcomes and dynamic allocation decisions. Allocation correctness is ensured at the structural level through the validation module, which checks output format and constraint satisfaction before finalizing each assignment; if validation fails, corrective feedback is issued and the LLM retries until a valid decision is produced. Ground truth optimal allocations are not available for this NP-hard combinatorial problem; instead, the RL baselines serve as strong performance references, with all methods evaluated under identical task sequences, robot capability models, and allocation constraints, ensuring that observed differences reflect purely algorithmic effectiveness.

4.4 Evaluation Metrics

To validate the framework’s effectiveness in managing the performance–fairness trade-off, this study employs two primary metrics:

Task Completion Rate (Success Rate): This measures the system’s ability to maximize execution feasibility. It is calculated as the ratio of successfully completed tasks to the total number of tasks requested. This metric serves as the primary benchmark for comparing the proposed LTAA framework against Q-learning and DQN baselines.

Workload Distribution Analysis: Instead of a single scalar balance score, we analyze the distribution of assigned tasks (W_r) across the heterogeneous robot team (Light, Medium, Heavy). This metric evaluates the framework’s internal logic in assigning tasks to capability-appropriate robots rather than overloading a single “generalist” agent. A superior distribution is defined as one that utilizes all robot types according to their specific feature strengths (Heavy vs. Dexterous) rather than allowing significant idleness in specialized units.

4.5 Experimental Protocol

Allocation decisions are generated sequentially to reflect the incremental nature of task assignment in construction projects. At each allocation step, the current task requirements, robot capability profiles, accumulated workload distribution, and project progress are provided to the allocation method. The proposed phase-adaptive approach dynamically adjusts decision priorities based on project progression, while baseline reinforcement learning methods operate with static allocation objectives.

All methods are evaluated using identical task sequences and robot capability configurations, to ensure fair comparison. Performance and workload balance metrics are aggregated across all tasks to evaluate overall allocation effectiveness and fairness.

5 Results

5.1 Task Completion Performance

Figure 2 compares task completion rates across the proposed phase-adaptive approach and reinforcement learning baselines. The proposed method achieves a task completion rate of 76%, performing competitively with Q-learning (73%) and Deep Q-Networks (77%).

LLM-based reasoning can match the performance of sophisticated optimization algorithms without the extensive training requirements inherent to RL. While DQN achieves a marginally higher success rate (1.03% difference), it requires significant hyperparameter tuning and training time. In contrast, LTAA achieves comparable performance in a zero-shot manner by leveraging semantic understanding of task features (Heavy, Dexterous, Careful) to align them with robot capabilities. This demonstrates that natural language reasoning is a viable substitute for mathematical optimization in construction tasks where training data may be scarce.

5.2 Workload Balance Analysis

Figure 3 illustrates the specific allocation distribution across the heterogeneous robot team, showing 22 tasks assigned to Light robots, 10 to Medium robots, and 4 to Heavy robots.

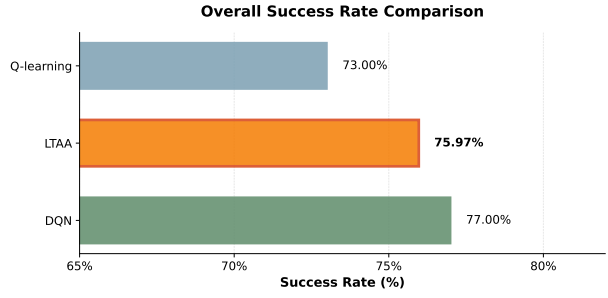


Figure 2. Task Completion Performance (Success Rate)

Robots Allocation across Light, Medium, and Heavy Tasks

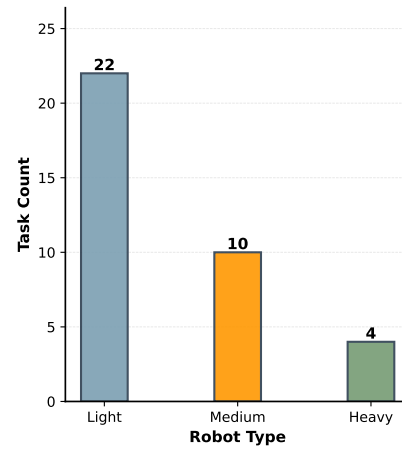


Figure 3. Workload Distribution

This distribution reveals the framework’s internal logic: rather than simply maximizing a global reward function (which often leads to “winner-take-all” behavior in RL where the most capable robot is overworked), LTAA distributes tasks according to specific capability matches.

Here the high assignment count for Light robots correlates with the prevalence of “Careful” and “Dexterous” tasks in the dataset, for which Light robots have higher success probabilities (0.90 and 0.80, respectively). Unlike standard greedy approaches that might assign a “Medium” generalist robot to everything, the LTAA framework correctly reserves the “Heavy” robot (4 tasks) strictly for force-intensive operations where it holds a comparative advantage (0.90 success rate). This confirms the system is successfully identifying capability niches rather than relying on a single “workhorse” agent.

5.3 Scenario-Based Performance Evaluation

To test robustness, the “Heavy Excels” scenario (Figure 4) evaluates the framework when one robot type has

significantly superior capabilities. In this configuration, LTAA achieves a 77.1% success rate while maintaining a balanced workload distribution (11, 12, and 13 tasks respectively).

The significance of this result lies in the Phase-Adaptive Strategy. In a static system, a “Heavy Excels” scenario would typically result in the Heavy robot receiving the vast majority of assignments due to its superior stats. However, Figure 4 demonstrates that LTAA successfully curbed this tendency. By shifting priorities toward equity in the middle and late phases, the framework forced the distribution of tasks to the Light (11) and Medium (12) robots despite the Heavy robot’s advantage. This confirms that the Phase-Adaptive logic effectively stops the strongest robot from monopolizing the workload. Instead of letting less capable robots sit idle, the system ensures the entire team remains active. This is a critical safety feature for construction, as it avoids overworking a single machine and reduces the risk of mechanical fatigue caused by relying on just one agent.

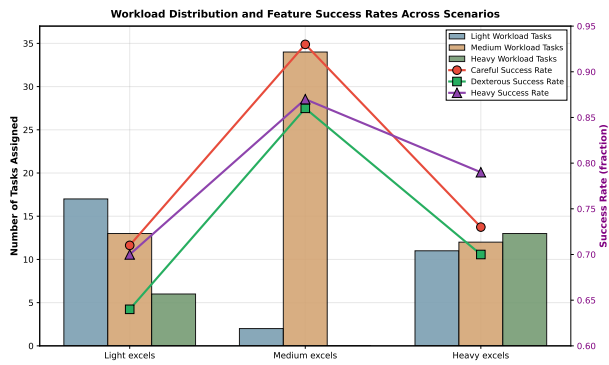


Figure 4. Scenario-Based Performance

6 Conclusions and Future Direction

This paper presented a phase-adaptive LLM-based task allocation approach for coordinating heterogeneous construction robot teams. Unlike static allocation strategies, the proposed method explicitly adjusts decision priorities across project phases, prioritizing task execution feasibility during early stages and progressively emphasizing workload equity as project progression advances. This phase-aware strategy reflects the temporal dynamics of construction operations, where early flexibility enables later fairness without compromising overall performance.

Through systematic benchmarking against reinforcement learning baselines, including Q-learning and Deep Q-Networks, the proposed approach demonstrated competitive task completion performance while achieving workload balance across robots. Scenario-based evaluation under pronounced capability asymmetry further

showed that phase-adaptive allocation effectively manages the performance–fairness trade-off in construction-relevant settings, avoiding persistent over-utilization of specialized robots.

These results indicate that LLM-based reasoning, when structured with phase-adaptive prioritization, can serve as a practical and interpretable alternative to learning-based task allocation methods for construction robotics. By operating without task-specific training or model fine-tuning, the approach offers flexibility for adapting to evolving construction requirements and heterogeneous robot capabilities.

Future work will focus on extending the phase-adaptive allocation strategy to handle incremental task arrivals and limited rescheduling within ongoing construction workflows. While the TEACH-derived task abstractions capture the multi-feature, sequential decision-making structure relevant to construction coordination. Evaluating the approach under additional robot capability configurations and larger task sets will further clarify its robustness and scalability. These extensions will support more realistic construction deployment scenarios while preserving the interpretability and adaptability demonstrated in this study.

References

- [1] Associated Builders and Contractors. 2024 construction workforce shortage tops half a million, 2024. Press release, Washington, DC. Available at: abc.org.
- [2] Tamima S. Elbashbisy and Islam El-adaway. Assessing the impact of skilled labor shortages on project cost performance. In *Proceedings of the Construction Research Congress*, pages 361–370, 2024.
- [3] Hongrui Yu, Vineet R. Kamat, Carol C. Menassa, William McGee, Y. Guo, and Hyun Lee. Mutual physical state-aware object handover in full-contact collaborative human–robot construction work. *Automation in Construction*, 150:104829, 2023.
- [4] Brian P. Gerkey and Maja J. Matarić. A formal analysis and taxonomy of task allocation in multi-robot systems. *The International Journal of Robotics Research*, 23(9):939–954, 2004.
- [5] Ernesto Nunes, Mitchell McIntire, and Maria Gini. Decentralized multi-robot allocation of tasks with temporal and precedence constraints. *Advanced Robotics*, 31(22):1193–1207, 2017.
- [6] S. Choudhury, J. K. Gupta, Mykel J. Kochenderfer, Dorsa Sadigh, and Jeannette Bohg. Dynamic multi-robot task allocation under uncertainty and tempo-

- ral constraints. *Autonomous Robots*, 46(1):231–247, 2022.
- [7] Dongmin Lee, SangHyun Lee, Neda Masoud, M. S. Krishnan, and Victor C. Li. Digital twin-driven deep reinforcement learning for adaptive task allocation in robotic construction. *Advanced Engineering Informatics*, 53:101710, 2022.
- [8] S. Kannan, V. Dhiman, L. Kloud, and David Hsu. Smart-llm: Smart multi-agent robot task planning using large language models. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1847–1854, 2024.
- [9] Peihan Li, Zijian An, Shams Abrar, and Lifeng Zhou. Large language models for multi-robot systems: A survey. *arXiv preprint arXiv:2502.03814*, 2025. URL <https://arxiv.org/abs/2502.03814>.
- [10] X. Ye, H. Guo, and Z. Luo. Two-stage task allocation for multiple construction robots using an improved genetic algorithm. *Automation in Construction*, 165:105583, 2024.
- [11] H. Chakraa, F. Guérin, E. Leclercq, and D. Lefebvre. Optimization techniques for multi-robot task allocation problems: Review on the state-of-the-art. *Robotics and Autonomous Systems*, 168:1–14, 2023.
- [12] L. Chen, W. L. Liu, and J. Zhong. An efficient multi-objective ant colony optimization for task allocation of heterogeneous unmanned aerial vehicles. *Journal of Computational Science*, 58:101545, 2022.
- [13] M. Deng, B. Fu, L. Li, and X. Wang. Integrating LLMs and digital twins for adaptive multi-robot task allocation in construction. *arXiv preprint arXiv:2506.18178*, 2025.
- [14] S. A. Prieto and B. García de Soto. Collaborative large language models for task allocation in construction robots. SSRN Working Paper No. 5097309, 2024.
- [15] K. Kim, P. Ghimire, and P.-C. Huang. Framework for LLM-enabled construction robot task planning: Knowledge base preparation and robot-LLM dialogue for interior wall painting. *Robotics*, 14(9):117, 2025.
- [16] Dong-Hyun Lee, Sheir Afzal Zaheer, Ji-Hyeong Han, Jong-Hwan Kim, and Eric Matson. Competency adjustment and workload balancing framework in multirobot task allocation. *International Journal of Advanced Robotic Systems*, 15(6):1729881418812960, 2018.
- [17] Fatma Faruq, David Parker, Bruno Lacerda, and Nick Hawes. Simultaneous task allocation and planning under uncertainty. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3559–3564. IEEE, 2018.
- [18] A. Padmakumar, J. Thomason, A. Shrivastava, P. Lange, A. Narayan-Chen, S. Gella, R. Piramuthu, G. Tur, and D. Hakkani-Tur. Teach: Task-driven embodied agents that chat. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 2017–2025, 2022.
- [19] Fabian Ranz, Vera Hummel, and Wilfried Sihm. Capability-based task allocation in human-robot collaboration. In *Procedia Manufacturing*, volume 9, pages 182–189. Elsevier, 2017.
- [20] Kehui Liu, Zixin Tang, Dong Wang, Zhigang Wang, Xuelong Li, and Bin Zhao. Coherent: Collaboration of heterogeneous multi-robot system with large language models. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 10208–10214. IEEE, 2025.
- [21] Glen Neville, Jiazhen Liu, Sonia Chernova, and Harish Ravichandar. Q-itags: Quality-optimized spatio-temporal heterogeneous task allocation with a time budget. *arXiv preprint arXiv:2404.07902*, 2024. URL <https://arxiv.org/abs/2404.07902>.
- [22] Bo Fu, William Smith, Denise M. Rizzo, Matthew Castanier, Maani Ghaffari, and Kira Barton. Robust task scheduling for heterogeneous robot teams under capability uncertainty. *IEEE Transactions on Robotics*, 39(2):1087–1105, 2022.