2026 European Conference on Computing in Construction
Corfu, Greece
July 12-15, 2026

# NEURO-SYMBOLIC REINFORCEMENT LEARNING FOR MULTI-CATEGORY CONSTRUCTION RELATIONSHIP WITH REASONING AND SEQUENCE PLANNING

Akarsth Kumar Singh and Shang Hsien Heish
National Taiwan University, Taipei, TAIWAN

## Abstract

Construction planning depends on tacit expertise that is rarely captured, causing recurring knowledge loss and weak standardization across projects. This paper presents a neuro-symbolic reinforcement learning framework for multi-category construction relationship reasoning that converts execution logic into explicit, domain-grounded constraints. It integrates synthetic WBS generation, LLM-based constrained reasoning, and Group Relative Policy Optimization (GRPO) with symbolic rewards to generate explainable relationship graphs at inference time. Results demonstrate stable learning, over 75% structural validity and reasoning quality across categories, bounded policy divergence, and reduced truncation, supporting more consistent, reusable, and auditable automated construction planning.

## Introduction

Construction planning in practice is rarely a standardized or transferable process. Instead, it relies heavily on planners' tacit knowledge, informal judgment, and accumulated site experience. This knowledge is seldom codified. When a project concludes, the sequencing logic and coordination strategies developed during execution leave with the planner, while organizations retain only static schedules that record outcomes rather than the reasoning behind them (Koo et al., 2007). As a result, planning knowledge is repeatedly lost, forcing each project to relearn previously acquired lessons and contributing to persistent inefficiencies in construction delivery (Wu et al., 2010; Amer et al., 2023).

This limitation is compounded by the fact that construction execution is governed by a wide range of construction execution constraints that extend far beyond what is explicitly represented in formal schedules. These constraints encompass interdependent structural, temporal, resource, cross-discipline, quality, and safety considerations that are typically managed implicitly through informal coordination rather than codified planning artifacts. When such coordination fails, the consequences include rework, schedule delays, cost overruns, and safety risks. The persistent gap between documented plans and the constraints that must be coordinated on site remains a primary source of inefficiency in construction projects (Pfitzner et al., 2024; Soman and Molina-Solana, 2022). Capturing and reasoning over these interacting constraints simultaneously imposes a cognitive burden that exceeds the limits of manual planning processes, particularly for large-scale projects.

In this context, artificial intelligence becomes a necessary enabler rather than a mere productivity enhancement. While other industries have adopted AI-driven planning and coordination, construction planning remains largely manual and experience-dependent (Rashid and Louis, 2019). Recent advances in large language models and reinforcement learning offer an opportunity to formalize and automate the reasoning processes that planners currently apply implicitly. However, for AI to deliver meaningful value in construction, it must move beyond isolated automation of scheduling logic and instead reason explicitly across interacting execution constraints in an explainable and transferable manner. Such relationship modeling is also foundational for downstream applications, including 4D BIM simulation, automated project controls, and delay analysis.

Existing research has addressed individual aspects of this problem. Reinforcement learning has been applied to optimize resource-constrained sequencing (Yao et al., 2024; Kedir et al., 2022), natural language processing models have inferred scheduling dependencies from historical data (Amer et al., 2023), BIM-driven methods have automated schedule generation from geometric models (Faghihi et al., 2014; Gan, 2022), and knowledge graphs have been used to structure construction process information (Pfitzner et al., 2024; Shan et al., 2025). Large language models have further shown promise in compliance checking and construction knowledge extraction (Yang and Zhang, 2024; Zhang et al., 2025). Despite these advances, existing approaches remain fragmented, typically addressing only a limited subset of construction relationships in isolation. None provides a unified framework for reasoning jointly across the full spectrum of execution constraints.

Based on this gap, three core challenges must be addressed to enable practical and transferable AI-driven construction planning.

**Challenge 1**: Multi-category reasoning. Construction relationships span multiple interacting categories, including scheduling logic, structural and physical constraints, resource requirements, temporal constraints, cross-discipline coordination, quality and inspection requirements, and safety and permitting rules. These categories are tightly coupled, as physical and regulatory constraints directly shape feasible sequencing and temporal dependencies. Effective reasoning therefore requires these relationships to be considered jointly rather than independently.

**Challenge 2**: Symbolic grounding. Construction relationships are governed by domain invariants derived from physical laws, material behavior, regulatory codes, and established construction methodologies. Predictions that violate load-path logic, omit mandatory inspections, or ignore curing requirements are not merely inaccurate but potentially unsafe. Learning-based models must therefore be grounded in explicit symbolic representations that enforce these constraints.

**Challenge 3**: Explainability and standardization. For AI-generated plans to be trusted, adopted, and reused across projects, each predicted relationship must be accompanied by a clear justification traceable to domain knowledge. Explicit reasoning enables standardization and knowledge transfer, directly addressing the chronic loss of planning knowledge observed in current practice.

To address these challenges, this research proposes a neuro-symbolic reinforcement learning framework for multi-category construction relationship reasoning. The framework integrates symbolic domain knowledge with learned reasoning policies through this research pipeline. First, a synthetic Work Breakdown Structure generation process produces structurally valid construction inputs at scale. Second, multi-category relationships are inferred using an ensemble of large language models constrained by formalized domain knowledge, generating reasoning-annotated dependency datasets across multiple execution constraint categories. Third, a 7B-parameter small language model is fine-tuned using Group Relative Policy Optimization with a composite reward function that encodes relationship correctness, directed acyclic graph validity, temporal plausibility, constraint compliance, and category coverage. Finally, the trained model generates explainable, multi-category relationship graphs for unseen inputs at inference time. Symbolic knowledge serves a dual role by grounding reasoning during data generation and providing verifiable reward signals during policy optimization.

The contributions are threefold. First, this research introduces a unified neuro-symbolic formulation for construction relationship reasoning, converting tacit planning knowledge into explicit, verifiable constraints rather than relying on symbolic rules or learning models alone. Second, it develops an LLM-driven data generation pipeline that produces reasoning-annotated construction relationship datasets with explicit justifications. Third, it provides a reproducible performance reference for multi-category dependency prediction using reinforcement learning with symbolic rewards, where LLM reasoning serves as the pre-optimization benchmark and supports explainable, standardized knowledge transfer across projects and organizations.

## Related Work

### AI for Construction Planning and Scheduling

Research on automated construction planning has progressed through several methodological paradigms. Early work emphasized constraint-based and rule-driven approaches using predefined methods, sequencing templates, and resource packages. For example, Wu et al. (2010) applied simulation and pattern-matching to generate schedules under explicit constraints, while Koo et al. (2007) formalized sequencing rationale through a constraint ontology for representing activity logic. Subsequent BIM-driven approaches leveraged geometric and structural information. Faghihi et al. (2014) employed genetic algorithms to derive structurally feasible sequences, and Gan (2022) introduced a graph-based model integrating IFC semantics and spatial topology for modular design generation. Despite their effectiveness, these methods typically address isolated relationship categories, such as scheduling or structural constraints, and depend on hand-coded rules that limit adaptability across projects.

Recent studies have shifted toward data-driven learning of construction sequencing. Amer et al. (2023) used transformer models to infer dependency constraints and generate lookahead plans, while Rashid and Louis (2019) demonstrated activity pattern recognition using deep learning with synthetic data. However, these approaches lack explicit symbolic grounding, failing to enforce domain invariants such as load transfer, curing constraints, and inspection requirements, thereby limiting reliability in safety-critical planning contexts.

### Reinforcement Learning in Construction

Reinforcement learning has emerged as a promising paradigm for optimizing construction scheduling decisions under uncertainty and resource constraints. Yao et al. (2024) proposed a deep reinforcement learning framework with valid action sampling and graph convolutional feature extraction, achieving reduced project durations compared to heuristic baselines. Kedir et al. (2022) combined reinforcement learning with agent-based modeling and graph embeddings to optimize activity sequencing under constrained resources. Soman and Molina-Solana (2022) applied Q-learning with linked-data-based constraint checking to automate lookahead schedule generation, illustrating the potential of reinforcement learning as a decision support tool for short-term planning. More recently, Yang et al. (2026) introduced the Construction Markov Decision Process framework, employing multi-agent proximal policy optimization to simulate crew-level decision-making across multiple construction stages with spatial dynamics.

Despite these advances, reinforcement learning applications in construction have largely focused on optimizing scheduling objectives such as project duration or resource utilization within narrowly scoped problem formulations. Symbolic domain knowledge is typically applied as hard constraints or pre-processing checks rather than being embedded directly into the reward structure. As a result, these models neither reason across multiple categories of construction relationships nor provide explicit, explainable justifications for predicted dependencies, limiting their usefulness for knowledge transfer and practitioner adoption.

### Knowledge Graphs, Ontologies, and Semantic Reasoning in Construction

Knowledge representation through ontologies and knowledge graphs has a long-standing role in construction informatics. Pauwels et al. (2017) enhanced the ifcOWL ontology with optimized geometric representations to improve semantic BIM data exchange. Zhang and El-Gohary (2017) integrated semantic NLP with logical reasoning for automated building code compliance and later extended this work using transformer-based models to align IFC entities and regulatory concepts within a knowledge-graph-enhanced classification framework Zhang and El-Gohary (2023). Ding et al. (2016) proposed an ontology-based framework for construction risk knowledge management, enabling semantic inference of risk interdependence in BIM environments. Collinge et al. (2022) developed a BIM-based safety ontology to support prevention-through-design by structuring tacit and explicit knowledge. More recently, Pfitzner et al. (2024) constructed knowledge graphs from continuous site image data, linking detected objects to spatiotemporal metrics for progress monitoring.

These studies demonstrate the value of structured knowledge representations for specific domains such as safety, compliance, risk management, and progress tracking. However, most knowledge graph–based systems remain static or query-driven and are not integrated with learning-based policy optimization. Moreover, they are typically domain-specific and do not provide a unified representation that jointly captures scheduling, structural, resource, temporal, cross-discipline, quality, and safety relationships required for automated construction planning.

### Large Language Models and Neuro-Symbolic Approaches in AEC

Large language models have recently been introduced to address knowledge-intensive tasks in the AEC domain. Yang and Zhang (2024) demonstrated prompt-based automation of building code information transformation, achieving high precision with limited training data. Zhang et al. (2025) proposed a JSON-based agent schema for LLM-driven building energy analysis, enabling reusable agent workflows. Shan et al. (2025) applied LLMs with in-context learning to automate knowledge graph construction for building health resilience using multi-modal data. More broadly, Pan et al. (2023) highlighted the potential of integrating LLMs, which encode parametric knowledge, with knowledge graphs, which represent explicit symbolic knowledge, to overcome limitations inherent in either paradigm alone.

Despite these promising developments, current LLM applications in construction remain predominantly zero-shot or few-shot and rely heavily on prompt engineering rather than systematic learning. Generated outputs are rarely validated against domain knowledge, often lack consistency across relationship categories, and provide no guarantees of physical or regulatory correctness. Furthermore, no prior work combines LLM-based symbolic reasoning for dataset generation with reinforcement learning–based policy optimization for construction dependency prediction. The framework proposed in this study addresses this gap through a neuro-symbolic pipeline in which formalized domain knowledge grounds LLM reasoning during data generation and simultaneously serves as a verifiable constraint signal during reinforcement learning optimization The reasoning-annotated outputs generated by the LLM during data construction serve as the reference against which the RL-trained model is evaluated, implicitly establishing prompted LLM reasoning as the performance reference point that policy optimization is designed to exceed.

## Problem Formulation

Given a Work Breakdown Structure (WBS) instance

$$W = \{h, p, l, O, T\}$$

where $h$ denotes the construction head (discipline), $p$ the work package, $l$ the location (building, floor, or zone), $O=\{o1,\ldots,om\}$ a set of structural or MEP objects, and $T=\{t1,\ldots,tn\}$ a set of construction tasks without explicit dependencies, the objective is to infer a multi-category construction relationship graph

$$G = \{(r_i, c_i, l_i, e_i)\}_{i=1}^{|G|}$$

Each relationship tuple consists of:

- $ri=(from,to,link\_type,lag)$, a directed dependency between two tasks or objects with an associated precedence type (FS, SS, FF, SF) and lag duration;
- $ci\in C$, the relationship category (scheduling, structural, resource, temporal, cross-discipline, quality, or safety);
- $li$, a natural-language justification traceable to domain knowledge;
- $ei\in[0,1]$, a confidence score.

Figure 1 illustrates a sample multi-category relationship graph for a column construction WBS scope, where nodes represent tasks and directed edges encode relationship tuples across structural, quality, temporal, and scheduling categories with associated confidence scores.

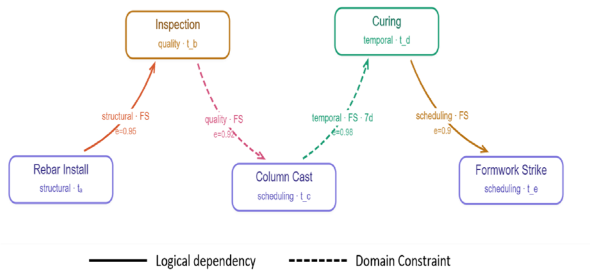


*Figure 1: Example multi-category construction relationship graph G inferred from a column construction WBS instance*

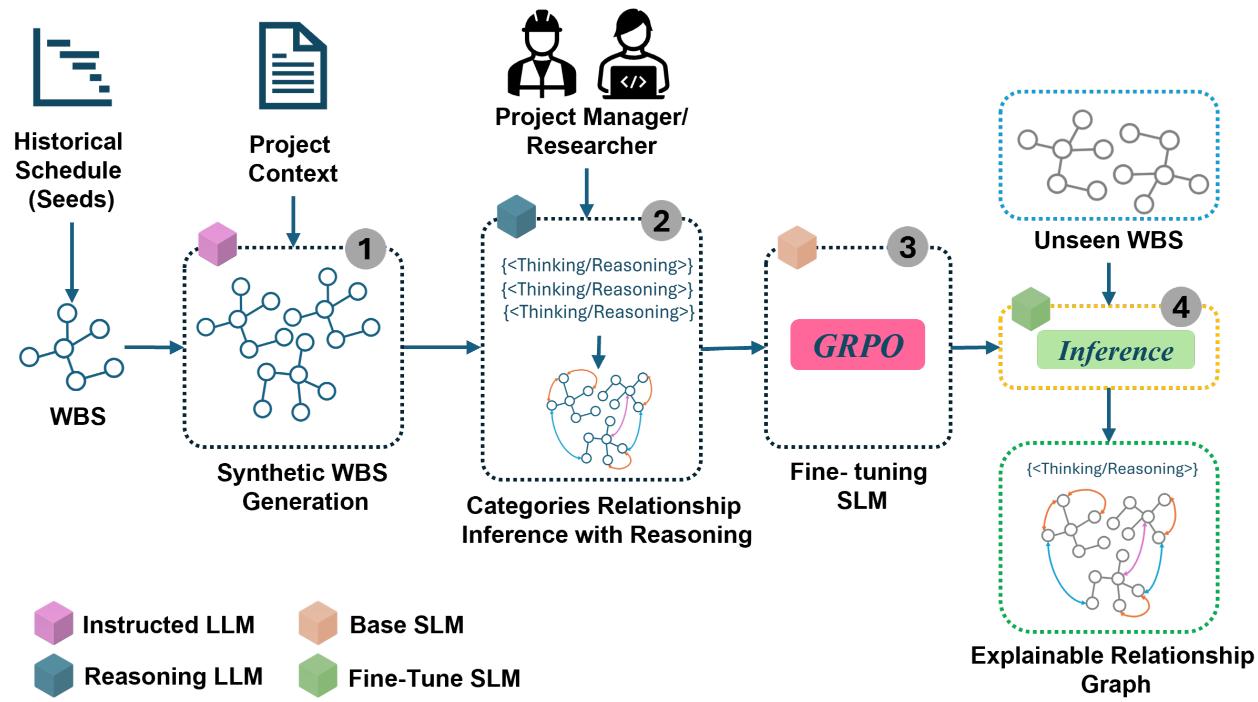


*Figure 2: Overview of Framework*

The predicted graph $G$ must satisfy three validity conditions:

(1) graph validity, such that scheduling relationships form a directed acyclic graph;

(2) physical feasibility, ensuring compliance with gravity-driven load transfer and construction logic; and

(3) domain compliance, meaning all applicable construction constraints are respected.

## Methodology

The proposed methodology integrates synthetic data generation, domain-constrained reasoning, and reinforcement learning to enable multi-category construction relationship inference. As illustrated in Figure 2, the framework begins with the generation of diverse Work Breakdown Structure representations reflecting realistic construction contexts without explicit dependency information. Construction relationships are then inferred category-by-category using large language model reasoning grounded in explicit domain constraints, generating directed edges between task nodes with accompanying explanatory traces. These outputs train a small language model through reinforcement learning, where symbolic constraints are embedded directly into the optimization objective as verifiable reward signals. During inference, the trained model generates explainable construction relationship graphs validated for scheduling acyclicity and physical compliance before being finalized across all execution categories.

### Domain Knowledge Representation

Construction knowledge that is typically implicit in planners' experience is formalized as a set of verifiable domain constraints expressed as predicates over the relationship graph G — for example, $\mathrm{supports}(o_i, o_j) \rightarrow \mathrm{precedes}(\mathrm{task}(o_i), \mathrm{task}(o_j))$ for structural load transfer, or $\mathrm{requires\_inspection}(t_i) \rightarrow \mathrm{precedes}(\mathrm{inspection}(t_i), t_j)$ for safety compliance. These constraints encode principles derived from structural mechanics, material behavior, construction methodology, trade coordination practice, and regulatory requirements. Each constraint maps to one or more relationship categories and can be automatically evaluated against a predicted graph, enabling both dataset validation and reward computation without manual annotation.

### Synthetic WBS Generation

Labeled multi-category construction data is scarce due to proprietary schedules and inconsistent documentation practices. To address this, synthetic WBS instances are generated using template-based augmentation grounded in real project schedules. Recurring WBS patterns are first extracted from reinforced-concrete building projects, capturing typical relationships among construction heads, work packages, locations, objects, and tasks. These patterns are then provided to an instruct-tuned LLM, which generates diverse yet domain-consistent WBS instances by varying contextual attributes such as floor levels, object types, and task labels. Each generated WBS is normalized into a unified schema containing tasks and objects but no explicit dependencies, ensuring that all relationships must be inferred by the model.

### Multi-Category Relationship Inference via LLM Reasoning

Each synthetic WBS instance is processed by a reasoning-oriented large language model to infer construction relationships across scheduling, structural, resource, temporal, cross-discipline, quality, and safety categories. Inference is performed in a category-aware manner, reasoning over one category at a time while maintaining global context consistency. The model receives structured WBS inputs with category-specific domain knowledge to generate structured relationship outputs while verifying assumptions. Category-specific schemas define representations such as precedence and lag (scheduling), load-transfer logic (structural), and clearance or inspection requirements (quality and safety). For each category, the model outputs relationships with justifications and confidence scores. Outputs are validated for logical and structural consistency and aggregated into a unified multi-category graph, retaining reasoning traces to support learning of reasoning quality alongside output correctness.

### Reinforcement Learning with Symbolic Rewards

The small language model is fine-tuned using Group Relative Policy Optimization (GRPO) to learn construction relationship reasoning policies. For each WBS input, eight candidate relationship graphs (K=8) are sampled and ranked relative to one another using a symbolic, multi-component reward function (weights reported in the Experiments section), and policy updates reinforce higher-quality reasoning trajectories.

$$R(G) = \alpha_1 R_{\mathrm{correct}} + \alpha_2 R_{\mathrm{format}} + \alpha_3 R_{\mathrm{structure}} + \alpha_4 R_{\mathrm{constraint}} + \alpha_5 R_{\mathrm{reason}}$$

$R_{correct}$ evaluates semantic agreement with reference relationship graphs generated during data construction.

$R_{format}$ enforces strict adherence to the required output schema.

$R_{structure}$ validates category-specific structural properties, such as acyclic scheduling graphs and physically feasible dependencies.

$R_{constraint}$ verifies consistency with construction domain knowledge relevant to each relationship category.

$R_{reason}$ rewards clear, internally consistent reasoning traces accompanying each prediction.

Each reward component is independently scaled to [0,1] prior to weighted aggregation, and GRPO advantages are computed as each candidate's reward minus the group mean, normalized by group standard deviation across the K=8 sampled graphs. This formulation allows multiple valid solutions while penalizing infeasible or poorly justified outputs. By embedding symbolic validation directly into the reward, the model internalizes domain-consistent reasoning rather than relying on post-hoc filtering. This design directly targets the known limitations of zero-shot LLM reasoning, inconsistency across relationship categories and lack of guaranteed domain compliance, as mentioned in the Related Work section.

### Inference and Explainable Outputs

At inference time, a WBS instance is provided to the fine-tuned SLM, which generates a multi-category relationship graph with concise natural-language justifications for each dependency, reflecting the domain principles that motivate relationships and enabling transparency and interpretability. Since each WBS instance is processed independently, multiple instances can be inferred in parallel, enabling full project-scale schedule generation with inference time scaling linearly with the number of instances. Automated post-processing verifies scheduling acyclicity, structural feasibility, and domain compliance, flagging rather than discarding violations to preserve explainability. The final output can be directly consumed by scheduling software, BIM-based 4D simulation tools, or project control systems, and because each dependency includes a traceable justification, the generated plans are inherently transferable across planners and projects, directly addressing the knowledge loss problem identified in the Introduction.

## Experiments

### Dataset Construction

The experimental dataset was constructed using the proposed Stage 1 and Stage 2 pipelines, with a real-world residential construction schedule as the seed source. The project includes 28 Work Breakdown Structure (WBS) instances comprising 125 activities across substructure, superstructure, and finishing works. To enhance structural diversity and generalization, these WBS patterns were extended into synthetic variants representing multiple structural systems, including masonry, reinforced concrete, composite, precast, steel, and timber structures, specifically to mitigate the residential sequencing and coordination bias introduced by the single-project seed source. The final dataset contains 619 unique construction tasks spanning disciplines, locations, and object types. Using the Stage 2 reasoning pipeline, multi-category relationships were inferred for each task across seven execution categories, yielding 4,333 relationship annotations. The dataset was designed to maintain balanced representation across categories, preventing bias during training.

### Reinforcement Learning Fine-Tuning

The policy model was initialized from Qwen3-4B-Base and fine-tuned using Group Relative Policy Optimization (GRPO). For each WBS input, eight candidate relationship graphs were sampled per training step and ranked using a symbolic, multi-component reward function. The reward combined output correctness (40%), format compliance (15%), output structure validity (25%), rule application consistency (10%), and reasoning quality (10%), ensuring balanced supervision across semantic accuracy, structural validity, and explanatory coherence. Training was conducted for five epochs with a batch size of 8 and a learning rate of $1\times10^{-4}$, using a cosine learning-rate scheduler with warm-up to improve training stability. All experiments were performed on a workstation equipped with an NVIDIA RTX 6000 Ada GPU (48 GB VRAM); the 4B-parameter model is additionally compatible with deployment on GPUs with 24 GB VRAM using standard quantization, broadening accessibility beyond high-end research hardware.

## Results and Discussion

### Training-Time Analysis (GRPO Optimization Dynamics)

Training results indicate that the proposed neuro-symbolic GRPO framework converges in a stable and controlled manner. As shown in Figure 3, improvements in total reward are primarily driven by sustained gains in output correctness and rule-consistent reasoning, indicating that the policy progressively learns semantically valid construction relationships rather than merely satisfying structural or formatting constraints. Rule application improves steadily throughout training, confirming increasing internalization of symbolic domain knowledge, while format compliance and structural validity saturate early and function mainly as stabilizing constraints.

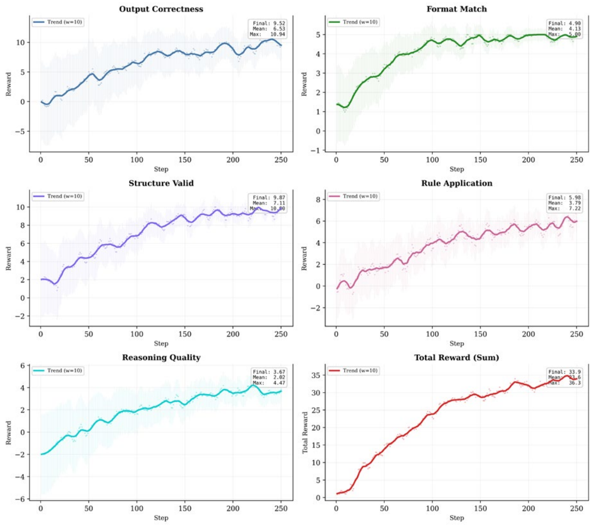

*Figure 3: Evolution of GRPO reward components during training*

Completion-length behavior further explains the improved optimization dynamics. As shown in Figure 4, increasing the maximum completion length reduces truncation, with the mean length stabilizing at approximately 1,370 tokens per WBS instance and an average inference time of 25–30 seconds per instance on the described hardware, indicating a predictable and bounded cost suitable for practical planning workflows. The truncation ratio decreases to 13.5%, preventing partial reasoning traces from corrupting reward computation and enabling more reliable evaluation of correctness and reasoning-related rewards.

Policy stability during GRPO optimization is supported by the KL divergence profile shown in Figure 5. After an initial adjustment phase, divergence remains bounded around a mean of approximately 0.22, with no upward drift over the training horizon, indicating controlled policy updates and regulated exploration. Overall, these results demonstrate reliable reward optimization,

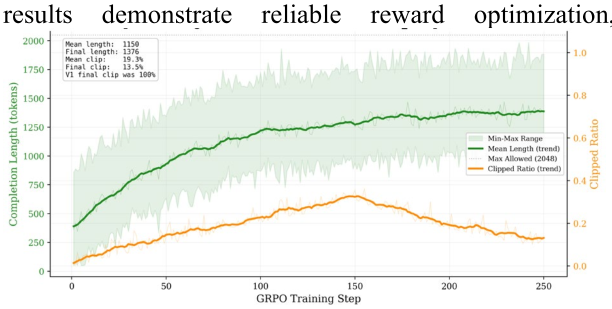

*Figure 4: Completion length and truncation behavior*

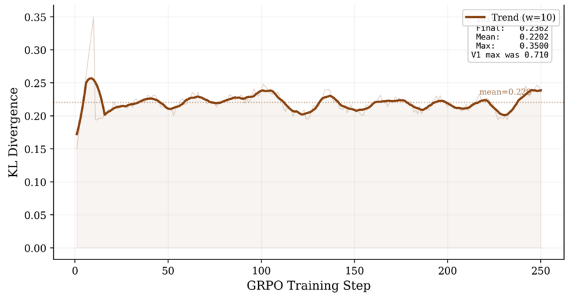

*Figure 5: KL divergence during GRPO optimization*

mitigation of truncation-induced noise, and stable learning dynamics, providing a robust foundation for inference-time evaluation on unseen construction planning instances.

**Inference-Time Performance and Category-wise Analysis**

Inference-time evaluation on 100 unseen WBS instances demonstrates that the trained model generalizes consistently across all reward components. As shown in Figure 6, the distributions of output correctness, structural validity, and format compliance are strongly concentrated near their upper bounds, indicating stable and reliable prediction behavior. Reasoning quality also exhibits a tight distribution around high scores, confirming that the model maintains coherent explanatory traces beyond training data. Rule application shows greater variance, reflecting the inherent difficulty of enforcing complex, context-dependent construction constraints during inference.

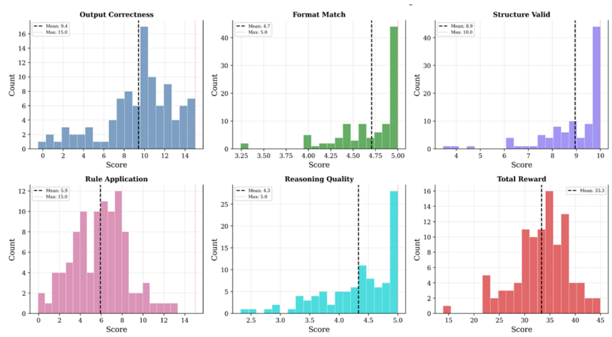

*Figure 6: Inference-time reward distribution*

A category-wise breakdown of inference rewards further highlights the multi-category robustness of the proposed framework. As illustrated in Figure 7, structural validity, format compliance, and reasoning quality consistently exceed the 75% performance threshold across all construction categories, including scheduling, structural, temporal, cross-discipline, quality, and safety. Output correctness varies more across categories, with lower scores observed for resource and cross-discipline reasoning, which require more implicit coordination knowledge. Nevertheless, safety and quality categories achieve the highest overall performance, reflecting the effectiveness of symbolic constraints in regulating critical execution logic. Overall, these results confirm that the

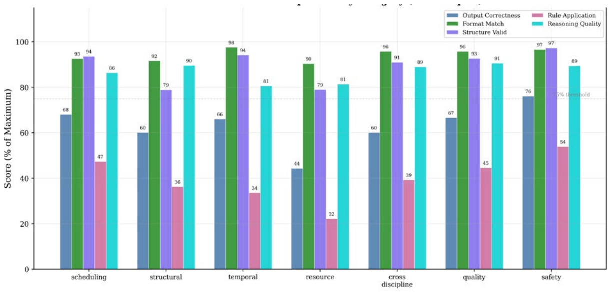


*Figure 7: Category-wise inference performance*

proposed neuro-symbolic GRPO framework produces explainable, structurally valid, and category-consistent construction relationship graphs at inference time, supporting its suitability for deployment in practical construction planning and downstream project control applications.

### Inference and Interpretability

Figure 8 shows an example of inference output for a foundation excavation task. From a dependency-free WBS input, the model generates coherent multi-category relationships, including scheduling precedence, temporal constraints, resource requirements, cross-discipline coordination, and safety considerations. The inferred relationships are context-aware, category-consistent, and aligned with practical construction logic. Non-applicable categories are omitted, indicating controlled reasoning rather than template-based generation. This example highlights the model's ability to produce interpretable and deployment-ready construction relationship graphs at inference time.

## Conclusions

This study presented a neuro-symbolic reinforcement learning framework for multi-category construction relationship reasoning, addressing key limitations in planning practice, namely the loss of tacit knowledge and the absence of standardized, explainable representations of execution constraints. Unlike prior approaches focused on isolated scheduling logic or post-hoc validation, the framework embeds symbolic domain knowledge into both data generation and policy optimization, enabling planning-time reasoning across interacting constraints.

Construction relationship inference is formulated as a structured graph prediction problem spanning scheduling, structural, temporal, resource, cross-discipline, quality, and safety categories. To mitigate limited labeled data, a synthetic Work Breakdown Structure generation pipeline is combined with reasoning-guided annotation using large language models constrained by domain knowledge. A small language model is subsequently fine-tuned using Group Relative Policy Optimization with a composite symbolic reward enforcing structural validity, temporal plausibility, constraint compliance, and reasoning quality, shifting learning from output memorization to transferable reasoning policies.

Experimental results indicate stable training dynamics, characterized by controlled policy updates, reduced truncation effects, and consistent reward convergence. Inference on unseen WBS inputs demonstrates robust generalization across all relationship categories, producing structurally valid and explainable dependency graphs. Performance is strong in structural and safety-critical relationships, while higher variability in resource and cross-discipline reasoning reflects the complexity of implicit coordination tasks. Each predicted relationship includes an explicit justification traceable to domain knowledge, supporting transparency and planner trust.

**System Prompt**

Analyze the construction task and identify scheduling relationships.

Output format: Array of scheduling_relationships with:
- from_task, to_task: Task names
- link_type: FS (Finish-to-Start), SS (Start-to-Start), FF (Finish-to-Finish), SF (Start-to-Finish)
- lag_days: Numeric lag in days
- reason: Justification citing relevant domain rules
- is_driving: Boolean indicating if this is on critical path
- total_float: Float days available

Output your analysis in this format:
<start_working_out>
[Your reasoning about the task, mentioning relevant domain rules]
<end_working_out>

<OUTPUT>
[JSON array of scheduling relationships]
</OUTPUT>

**Inference Input (User Prompt)**

Country: India, Project Type: Residential Building, Structure Type: RCC Frame, Work Package: Sub Structure (STRUCTURAL), Location: Foundation Area, Object: Foundation Footings, Task: Excavation Work

Analyze this task from a SCHEDULING perspective.

**Model Reasoning (Intermediate)**

<start_working_out>
Excavation work is a critical foundation activity and cannot begin until soil conditions and access safety are confirmed.
Per R-007 and R-010, temporary works such as shoring must be completed before excavation starts.
Soil assessment directly governs excavation feasibility and is therefore a driving predecessor.
Once excavation is underway, some verification activities such as subgrade testing can proceed in parallel in completed zones, forming a Start-to-Start relationship.
However, foundation formwork requires a short stabilization period after excavation completion, consistent with R-001 and R-009, introducing a Finish-to-Start dependency with lag.
Only preparatory and formwork-start dependencies are critical; intermediate testing activities retain float.
<end_working_out>

**Final Output (Scheduling Relationships)**

<OUTPUT>
[{ "from_task": "Soil Condition Assessment", "to_task": "Excavation Work","link_type": "FS","lag_days": 0, "reason": "Soil bearing capacity and type must be verified prior to excavation to ensure design compliance and safety.", "is_driving": true, "total_float": 0 },
{ "from_task": "Temporary Works Completion", "to_task": "Excavation Work", "link_type": "FS", "lag_days": 0, "reason": "Shoring and temporary stabilization systems must be in place before excavation to prevent collapse and allow safe access.", "is_driving": true, "total_float": 0 },
{ "from_task": "Excavation Work", "to_task": "Subgrade Compaction Testing", "link_type": "SS", "lag_days": 0, "reason": "Compaction and density testing can start in completed zones while excavation continues elsewhere.", "is_driving": false, "total_float": 2 },
{ "from_task": "Excavation Work", "to_task": "Foundation Formwork Start", "link_type": "FS", "lag_days": 1, "reason": "A minimum stabilization and preparation period is required after excavation before formwork erection begins.", "is_driving": true, "total_float": 0 }
]
</OUTPUT>

*Figure 8: Inference Output Showing Explainable Multi-Category Construction Relationships*

The study is based on a residential project dataset augmented with synthetic WBS generation, enabling scalable training but introducing bias toward residential construction logic. Although methodologically general, the learned policy may not transfer directly to other typologies, such as suspension bridges, tunnels, or petrochemical plants, where execution constraints differ significantly and the domain constraint layer must be adapted to project-specific physics, regulations, and construction methods. Future work will extend the constraint framework to additional project types, incorporate typology-specific constraints and uncertainty-aware reasoning, and validate on diverse real-world datasets to assess cross-domain transferability. Overall, this work demonstrates a practical pathway toward explainable and transferable AI-driven construction planning that respects the domain-specific nature of construction knowledge while enabling reuse across projects and organizations.

## Acknowledgments

The authors thank the National Science and Technology Council, Taiwan, for supporting this research through research grant Nos. NSTC 114-2119-M-002-034.

## References

Amer, F., Jung, Y. & Golparvar-Fard, M. (2023) Construction schedule augmentation with implicit dependency constraints and automated generation of lookahead plan revisions. Automation in Construction, 152, 104896.

Collinge, W.H., Farghaly, K., Mosleh, M.H., Manu, P., Cheung, C.M. & Osorio-Sandoval, C.A. (2022) BIM-based construction safety risk library. Automation in Construction, 141, 104391.

Ding, L.Y., Zhong, B.T., Wu, S. & Luo, H.B. (2016) Construction risk knowledge management in BIM using ontology and semantic web technology. Safety Science, 87, pp.202–213.

Faghihi, V., Reinschmidt, K.F. & Kang, J.H. (2014) Construction scheduling using genetic algorithm based on building information model. Expert Systems with Applications, 41(16), pp.7565–7578.

Gan, V.J.L. (2022) BIM-based graph data model for automatic generative design of modular buildings. Automation in Construction, 134, 104062.

Kedir, N.S., Somi, S., Fayek, A.R. & Nguyen, P.H.D. (2022) Hybridization of reinforcement learning and agent-based modeling to optimize construction planning and scheduling. Automation in Construction, 142,104498.

Koo, B., Fischer, M. & Kunz, J. (2007) Formalization of construction sequencing rationale and classification mechanism to support rapid generation of sequencing alternatives. Journal of Computing in Civil Engineering, 21(6), pp.423–433.

Pan, J.Z., Razniewski, S., Kalo, J.-C., Singhania, S., Chen, J., Dietze, S., Jabeen, H., Omeliyanenko, J., Zhang, W., Lissandrini, M., Biswas, R., de Melo, G., Bonifati, A., Vakaj, E., Dragoni, M. & Graux, D. (2023) Large language models and knowledge graphs: Opportunities and challenges. arXiv preprint, arXiv:2308.06374.

Pauwels, P., Krijnen, T., Terkaj, W. & Beetz, J. (2017) Enhancing the ifcOWL ontology with an alternative representation for geometric data. Automation in Construction, 80, pp.77–94.

Pfitzner, F., Braun, A. & Borrmann, A. (2024) From data to knowledge: Construction process analysis through continuous image capturing, object detection, and knowledge graph creation. Automation in Construction, 164, 105451.

Rashid, K.M. & Louis, J. (2019) Times-series data augmentation and deep learning for construction equipment activity recognition. Advanced Engineering Informatics, 42, 100944.

Shan, T., Zhang, F., Chan, A.P.C., Zhu, S. & Li, K. (2025) Large language models-empowered automatic knowledge graph development based on multi-modal data for building health resilience. Advanced Engineering Informatics, 68, 103655.

Singh, A.K. & Hsieh, S.H. (2026) Multi-LLM-based augmentation and synthetic data generation of construction schedules and task descriptions with SLM-as-a-judge assessment. Advanced Engineering Informatics, 69, 103825.

Soman, R.K. & Molina-Solana, M. (2022) Automating look-ahead schedule generation for construction using linked-data based constraint checking and reinforcement learning. Automation in Construction, 134, 104069.

Wu, I.C., Borrmann, A., Beißert, U., König, M. & Rank, E. (2010) Bridge construction schedule generation with pattern-based construction methods and constraint-based simulation. Advanced Engineering Informatics, 24(4), pp.379–388.

Yang, B., Liu, B., Han, Y., Meng, X., Wang, Y., Yang, H. & Xiao, J. (2026) Multi-agent reinforcement learning-enhanced decision-making of crew agents in floor construction processes. Advanced Engineering Informatics, 69, 103996.

Yang, F. & Zhang, J. (2024) Prompt-based automation of building code information transformation for compliance checking. Automation in Construction, 168, 105817.

Yao, Y., Tam, V.W.Y., Wang, J., Le, K.N. & Butera, A. (2024) Automated construction scheduling using deep reinforcement learning with valid action sampling. Automation in Construction, 166, 105622.

Zhang, J. & El-Gohary, N.M. (2017) Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. Automation in Construction, 73, pp.45–57.

Zhang, L., Fu, X., Li, Y. & Chen, J. (2025) Large language model-based agent schema and library for automated building energy analysis and modeling. Automation in Construction, 176, 106244.

Zhang, R. & El-Gohary, N. (2023) Transformer-based approach for automated context-aware IFC-regulation semantic information alignment. Automation in Construction, 145, 104540.