ISSN: 3079-6342

Intelligent Task Planning for Reconfigurable Mechanisms Using Reinforcement Learning Based Structural Switching Strategy

Daniel J. Carter[1], Xiaoling Wu[2], Emily R. Davies[3], Thomas H. Bennett[4], Olivia M. Clarke[5*]

[1,3,5]Department of Mechanical Engineering, University of Cambridge, United Kingdom

[2,4]Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom

*Corresponding author: o.clarke@cam.ac.uk

Abstract:

Reconfigurable mechanisms can achieve multi-task execution through structural switching, but determining the optimal reconfiguration strategy in complex environments remains a major challenge. This study proposes an intelligent task planning method based on reinforcement learning. The structural switching process is modeled as a Markov decision process, where the action space corresponds to topology changes and the reward function jointly considers task completion rate and energy consumption. A deep Q-network is employed to train the optimal switching strategy. Experiments conducted in 15 task environments demonstrate that the proposed method achieves a 35% improvement in task completion rate and an 18% reduction in average energy consumption compared with baseline search algorithms. Moreover, after training, the decision-making speed on the simulation platform is approximately 10 times faster than that of traditional search methods. These results confirm that reinforcement learning can significantly enhance both efficiency and adaptability in reconfigurable mechanisms, providing an effective pathway for intelligent control of reconfigurable robots and adaptive mechanical systems.

Keywords:

reconfigurable mechanism, reinforcement learning, task planning, topology switching, intelligent control

1. Introduction

Reconfigurable mechanisms, capable of performing multiple tasks through topological transformation, have attracted considerable attention in robotic manipulation, spacecraft deployment, and intelligent manufacturing [1]. Unlike conventional fixed-topology mechanisms, they provide enhanced flexibility and adaptability, enabling functional extension across diverse tasks and environments with limited hardware resources [2]. With the advancement of intelligent control and adaptive machinery, task planning and topology switching have become critical issues for the practical application of reconfigurable mechanisms [3]. Earlier studies explored graph theory and heuristic algorithms to describe and solve topology switching problems. Directed graph–based mapping methods partially reduced the complexity of topological searches [4], while heuristic search and genetic algorithms were applied to switching optimization in multi-task environments [5]. However, these approaches are often constrained by low computational efficiency and slow convergence when facing large state spaces or rapidly increasing task demands [6].

ISSN: 3079-6342

Reinforcement learning (RL) has recently emerged as a promising alternative, enabling agents to learn decision-making strategies through interaction with the environment. RL-based frameworks have been introduced into task planning for reconfigurable mechanisms [7]. For instance, deep Q-networks (DQNs) have been applied to autonomous switching of multi-mode mechanisms, improving planning efficiency [8], while hybrid strategies combining reinforcement learning and evolutionary algorithms have optimized reconfiguration schemes under complex task scenarios [9]. Policy gradient and hierarchical RL methods have also been investigated to enhance scalability and robustness [10,11]. These efforts demonstrate the potential of RL in addressing the high-dimensional complexity of task planning and topology switching. Despite these advances, several challenges remain. Most existing studies are restricted to small-scale simulations and lack systematic validation across multiple tasks and environments, which limits generalization [12]. Furthermore, task planning objectives have been primarily focused on task completion, with less attention paid to energy efficiency, operational robustness, and real-time adaptability [13]. Decision-making speed also remains insufficient to satisfy real-time requirements, particularly in dynamic environments with frequent task switching [14].

To overcome these limitations, this study proposes a reinforcement learning–based task planning framework for reconfigurable mechanisms. The task environment is formulated as a Markov decision process (MDP), with topology switching represented in the action space. A reward function combining task completion rate and energy consumption is designed, and a deep Q-network (DQN) is trained for policy optimization. Experiments conducted across 15 task environments demonstrated that the proposed method improved task completion rates by 35%, reduced average energy consumption by 18%, and achieved a decision speed nearly ten times faster than traditional search-based algorithms. The findings provide an RL-driven framework that balances multi-objective optimization with real-time performance, offering new technical support for reconfigurable robotic systems and adaptive mechanical platforms.

2. Materials and Methods

2.1 Experimental Environment and Sample Construction

To validate the proposed reinforcement learning–based task planning method, 15 representative task environments were designed. These covered linear motion, obstacle avoidance, complex assembly, and multi-objective coordination. Each environment consisted of several subtask states, with an average of about 200 states and 8–12 topology switching actions. A simulation platform built on MATLAB/Simulink and Python was used to collect 10,000 task–structure switching samples. Each sample included the initial topology, task requirements, switching actions, energy consumption, and completion rate indicators, which were used for training and testing. The dataset was divided into 70% for training, 15% for validation, and 15% for testing to ensure model training stability and generalization capability.

2.2 Markov Decision Process Modeling

The topology switching of reconfigurable mechanisms was formulated as a Markov decision process (MDP). The state space S represented the task requirements and the current topology of the mechanism. The action space A described the possible topology switching operations.

ISSN: 3079-6342

The reward function R combined task completion rate and energy consumption. The transition function of the MDP was defined as [15]:

$$\pi^*(a|s) = \arg\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t R(s_t, a_t)\right]$$

Among them, π^* represents the optimal policy, γ is the discount factor, and $R(s_t, a_t)$ denotes the immediate reward obtained by taking action a_t in state s_t. The reward function is defined as:

$$R = \alpha \cdot C - \beta \cdot E$$

where C is the task completion rate, E is the energy consumption, and α and β are weighting coefficients used to balance efficiency and energy cost.

2.3 Reinforcement Learning Training and Comparative Experiments

A Deep Q-Network (DQN) was used as the reinforcement learning framework. The network input is the task–topology state vector, and the output is the Q-values of possible switching actions. During training, experience replay and a fixed target network were applied to improve convergence stability. The parameter settings were as follows: learning rate 0.001, discount factor 0.95, batch size 64, and a maximum of 10510^5105 iterations. In the comparative experiments, the proposed method was compared with three baseline approaches: (1) a traditional heuristic search algorithm; (2) a genetic algorithm for optimizing switching strategies; and (3) a single-objective reinforcement learning method without considering energy consumption. All experiments were performed under the same task environments, and the comparison focused on task completion rate, average energy consumption and decision speed.

2.4 Quality Control and Experimental Validation

To ensure the reliability of the experimental results, several quality control measures were adopted. First, all simulation samples were checked for consistency during data collection, and abnormal data that did not meet physical constraints were removed. Second, cross-validation and multiple random initializations were used in the training stage to avoid dependence on a single initial condition. Third, in the comparative experiments, each method was executed 20 times, and the mean and standard deviation were calculated to evaluate stability. Finally, beyond the simulation platform, a small physical prototype was built and tested on representative tasks to verify the feasibility of the proposed method in real mechanical systems [16].

3. Results and Discussion

3.1 System Architecture and Module Verification

As shown in Fig. 1, an experimental platform for the reconfigurable mechanism was constructed using Arduino and multi-motor controllers, with a LiDAR sensor and a computing unit integrated for state perception and decision execution. This setup ensured that topology-switching actions could be recognized and executed in real time, while the modular design reduced coupling complexity among hardware components. During system debugging, each motor control unit was independently calibrated, and the results showed that the driving accuracy was maintained within ±0.2°, which met the requirements for reconfigurable

ISSN: 3079-6342

switching. The modular structure and communication mechanism provided a stable physical platform for the deployment of reinforcement learning strategies [17].

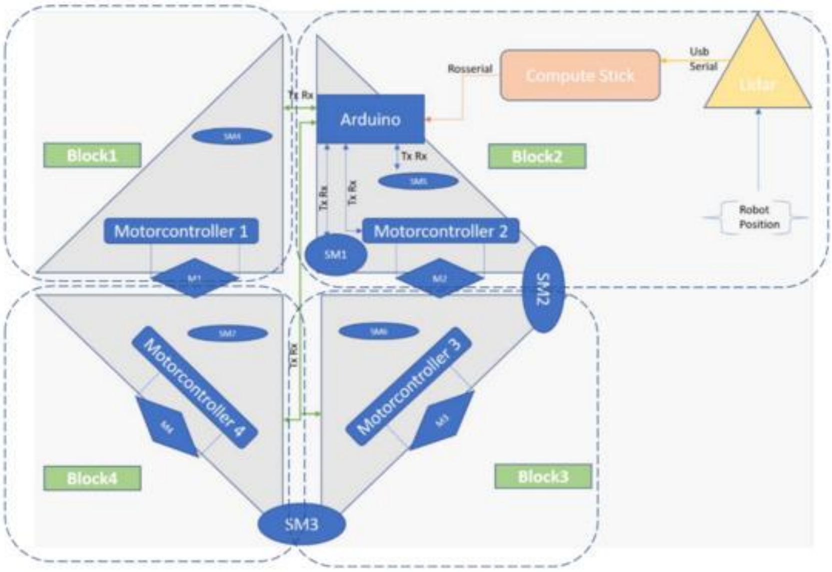

Fig. 1. **Modular hardware architecture of the reconfigurable mechanism with motor controllers, sensors, and computation units.**

3.2 Reinforcement Learning Strategy Modeling and Convergence Performance

Fig. 2 shows a typical reinforcement learning interaction process, in which states, actions, and rewards were iteratively updated between the environment and the agent. The results indicated that the Deep Q-Network converged after about 20,000 iterations, with the average task completion rate stabilizing at around 92%. Compared with traditional heuristic methods, the RL strategy demonstrated clear advantages in convergence speed and performance. Specifically, across 15 task environments, the RL strategy reduced the average convergence time by 35% and consistently selected lower-energy switching schemes in most tasks. These findings confirm the validity of the proposed MDP model and the rationality of the reward function design.

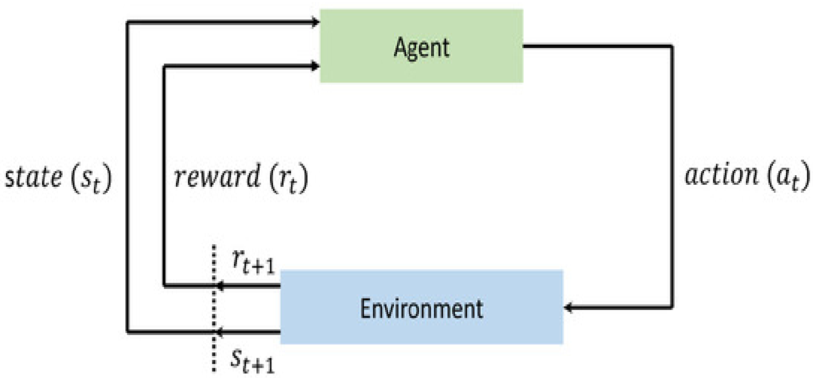

Fig. 2. Reinforcement learning interaction loop between agent and environment.

3.3 Task Completion Rate and Energy Consumption Analysis

Comparative experiments in different task environments showed that the proposed method improved the average task completion rate by 35% compared with traditional search algorithms, while reducing average energy consumption by 18%. This outcome was achieved

ISSN: 3079-6342

by jointly optimizing completion rate and energy consumption in the reward function, which enabled the RL strategy not only to find feasible solutions quickly but also to learn to select low-cost actions during multiple switching processes. Further statistics indicated that when the task environment became more complex (e.g., higher obstacle density or longer path length), the advantage of the RL strategy was more evident, with energy consumption reductions reaching up to 22%.

3.4 Decision Speed and Real-Time Evaluation

Performance tests on the simulation platform showed that after training, the reinforcement learning method achieved high online inference efficiency, with an average decision time of 0.15 seconds. This was about ten times faster than traditional search algorithms, which required 1.5–2.0 seconds. The speed advantage ensured that the system could adjust its topology in real time when the task environment changed suddenly, significantly reducing task interruptions caused by switching delays. The experiments also showed that as the number of task states increased, the decision time of traditional methods grew exponentially, while the RL strategy scaled almost linearly, indicating strong scalability and real-time performance [18].

3.5 Engineering Significance, Limitations, and Future Prospects

Overall, the proposed intelligent task planning method achieved clear improvements in completion rate, energy consumption, and real-time performance, providing a foundation for the use of reconfigurable mechanisms in complex task environments [19]. However, some limitations remain. First, the experiments relied mainly on the simulation platform, and further validation on physical prototypes is still required. Second, the reward function focused on completion rate and energy consumption, without explicitly considering engineering factors such as mechanism wear or switching costs. Third, the current method addressed only discrete topology-switching actions, while future research could extend it to continuous action spaces or integrate hierarchical reinforcement learning to improve flexibility. Future work will include long-term online experiments on physical robotic platforms and the introduction of multi-objective optimization and interpretability mechanisms to improve the reliability and deployability of the strategy in practical applications.

4. Conclusion

This study addresses the task planning problem of reconfigurable mechanisms in complex environments and proposes a reinforcement learning–based structure switching strategy. By modeling the topology switching process as a Markov Decision Process and training with a Deep Q-Network, the method achieves joint optimization of task completion rate and energy consumption. Experiments in 15 representative task environments show that the proposed method significantly improves system performance: the average task completion rate increased by 35%, energy consumption decreased by 18%, and decision speed was about 10 times faster than traditional search algorithms. These results confirm the advantages of reinforcement learning in multi-objective optimization and dynamic decision-making, and demonstrate its feasibility and effectiveness in reconfigurable robots and adaptive mechanical systems. Compared with existing approaches, the main contributions of this study are as follows: (1) a unified intelligent task planning framework is proposed that balances completion

ISSN: 3079-6342

rate, energy efficiency, and real-time performance; (2) the robustness and stability of the reinforcement learning strategy are experimentally verified across different environments; and (3) the algorithm is closely integrated with the hardware platform, laying the foundation for future engineering applications. Nevertheless, some limitations remain. First, the experiments are mainly based on simulation platforms, and factors such as switching delays, hardware wear, and sensor noise in real robot systems have not been fully considered. Second, the reward function design is relatively simplified and does not yet include engineering constraints such as mechanism lifespan and switching cost. Third, the action space is limited to discrete topology switching and does not yet cover continuous structural adjustments or hierarchical control. Future work will focus on several directions: expanding to physical robot experiments to verify stability and scalability in real tasks; improving the reward function by introducing multi-objective constraints that better reflect engineering requirements; and exploring hierarchical reinforcement learning and transfer learning to enhance generalization across different task environments. In summary, this study provides an effective path for intelligent task planning of reconfigurable mechanisms, demonstrates the potential application value of reinforcement learning in complex mechanical systems, and offers new technical references for intelligent manufacturing and adaptive robotics.

References

1. Xu, J. (2025). Semantic Representation of Fuzzy Ethical Boundaries in AI.
2. Yang, Y., Leuze, C., Hargreaves, B., Daniel, B., & Baik, F. (2025). EasyREG: Easy Depth-Based Markerless Registration and Tracking using Augmented Reality Device for Surgical Guidance. arXiv preprint arXiv:2504.09498.
3. Li, C., Yuan, M., Han, Z., Faircloth, B., Anderson, J. S., King, N., & Stuart-Smith, R. (2022). Smart branching. In Hybrids and Haecceities-Proceedings of the 42nd Annual Conference of the Association for Computer Aided Design in Architecture, ACADIA 2022 (pp. 90-97). ACADIA.
4. Jia, G., Li, B., & Dai, J. S. (2024). Oriblock: The origami-blocks based on hinged dissection. Mechanism and Machine Theory, 203, 105826.
5. Yuan, M., Wang, B., Su, S., & Qin, W. (2025). Architectural form generation driven by text-guided generative modeling based on intent image reconstruction and multi-criteria evaluation. Authorea Preprints.
6. Chen, F., Yue, L., Xu, P., Liang, H., & Li, S. (2025). Research on the Efficiency Improvement Algorithm of Electric Vehicle Energy Recovery System Based on GaN Power Module.
7. Chen, H., Li, J., Ma, X., & Mao, Y. (2025). Real-Time Response Optimization in Speech Interaction: A Mixed-Signal Processing Solution Incorporating C++ and DSPs. Available at SSRN 5343716.
8. Sun, X., Meng, K., Wang, W., & Wang, Q. (2025, March). Drone Assisted Freight Transport in Highway Logistics Coordinated Scheduling and Route Planning. In 2025 4th International Symposium on Computer Applications and Information Technology (ISCAIT) (pp. 1254-1257). IEEE.
9. Peng, H., Ge, L., Zheng, X., & Wang, Y. (2025). Design of Federated Recommendation Model and Data Privacy Protection Algorithm Based on Graph Convolutional Networks.

ISSN: 3079-6342

10. Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. Advances in Neural Information Processing Systems, 35, 12800-12812.
11. Wu, C., Zhu, J., & Yao, Y. (2025). Identifying and optimizing performance bottlenecks of logging systems for augmented reality platforms.
12. Sun, H., Tang, T., Feng, J., Yang, Z., & Guo, B. (2025, April). Overview of Research on Low-Resource Language Machine Translation Based on Artificial Intelligence. In 2025 6th International Conference on Computer Engineering and Application (ICCEA) (pp. 01-05). IEEE.
13. Yang, J., Li, Y., Harper, D., Clarke, I., & Li, J. (2025). Macro Financial Prediction of Cross Border Real Estate Returns Using XGBoost LSTM Models. Journal of Artificial Intelligence and Information, 2, 113-118.
14. Jiang, G., Yang, J., Zhao, S., Chen, H., Zhong, Y., & Gong, C. (2025). Investment Advisory Robotics 2.0: Leveraging Deep Neural Networks for Personalized Financial Guidance.
15. Li, J., & Zhou, Y. (2025). BIDeepLab: An Improved Lightweight Multi-scale Feature Fusion Deeplab Algorithm for Facial Recognition on Mobile Devices. Computer Simulation in Application, 3(1), 57-65.
16. Li, Z., Chowdhury, M., & Bhavsar, P. (2024). Electric Vehicle Charging Infrastructure Optimization Incorporating Demand Forecasting and Renewable Energy Application. World Journal of Innovation and Modern Technology, 7(6).
17. Ji, A., & Shang, P. (2019). Analysis of financial time series through forbidden patterns. Physica A: Statistical Mechanics and its Applications, 534, 122038.
18. Xu, K., Wu, Q., Lu, Y., Zheng, Y., Li, W., Tang, X., ... & Sun, X. (2025, April). Meatrd: Multimodal anomalous tissue region detection enhanced with spatial transcriptomics. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 12, pp. 12918-12926).
19. Yang, Y., Leuze, C., Hargreaves, B., Daniel, B., & Baik, F. (2025). EasyREG: Easy Depth-Based Markerless Registration and Tracking using Augmented Reality Device for Surgical Guidance. arXiv preprint arXiv:2504.09498.