TARAS SHEVCHENKO NATIONAL UNIVERSITY OF KYIV
(FACULTY OF INFORMATION TECHNOLOGY,
FACULTY OF COMPUTER SCIENCE AND CYBERNETICS)
NATIONAL TECHNICAL UNIVERSITY OF UKRAINE
“IGOR SIKORSKY KYIV POLYTECHNIC INSTITUTE”
VIKTOR GLUSHKOV INSTITUTE OF CYBERNETICS OF THE NAS OF UKRAINE
INSTITUTE FOR INFORMATION RECORDING OF THE NAS OF UKRAINE
INSTITUTE OF SOFTWARE SYSTEMS OF THE NAS OF UKRAINE
THE COUNCIL OF YOUNG SCIENTISTS OF THE FACULTY OF COMPUTER SCIENCE AND CYBERNETICS AND THE FACULTY OF INFORMATION TECHNOLOGY OF TARAS SHEVCHENKO NATIONAL UNIVERSITY OF KYIV
PUBLIC ASSOCIATION «COGNITIVE RESEARCH»

XI INTERNATIONAL CONFERENCE

Information Technology and Implementation (Satellite)

21 November, 2025

Conference Proceedings

Kyiv

2025

Vadym Ziuziun
PhD, Associate Professor of the Department of Technologies Management
Andrii Utvenko
PhD student of the Department of Technologies Management
Taras Shevchenko National University of Kyiv

APPLICATION OF REINFORCEMENT LEARNING TO OPTIMIZE INPUT PARAMETERS IN ORDER TO ACHIEVE PREDEFINED PERFORMANCE METRICS DURING BENCHMARKING OF A HARDWARE–SOFTWARE SYSTEM

Abstract. The growing complexity and versatility of contemporary software-hardware systems demand advanced methods for parameter selection and configuration optimization, particularly in the context of benchmarking tasks. Traditional manual approaches are often inefficient and may fail to deliver optimal results in dynamic and multifactorial environments. Reinforcement learning emerges as a promising paradigm for building autonomous agents capable of intelligently searching for and selecting input parameters, progressively improving system performance through iterative interaction with the environment.

Keywords: reinforcement learning, hardware–software system, optimization, benchmarking, adaptability, autonomy, hyperparameter optimization.

Modern hardware–software systems feature high levels of technical complexity, incorporating numerous functional subsystems and software modules that must operate efficiently in a diverse range of conditions. One of the central challenges in their testing and deployment is the effective selection of input parameters to achieve desired indicators such as performance, energy efficiency, or stability. This challenge is further amplified by the variability of real-world scenarios, the dynamic nature of operational environments, and the multifactorial influence of both internal and external factors.

The application of reinforcement learning (RL) offers new opportunities for the automation and intelligent adaptation of this process. RL agents learn iteratively by interacting with the system: by adjusting configuration parameters, they observe changes in target metrics, receive rewards or penalties according to outcomes, and gradually formulate optimal decision-making policies. Unlike traditional approaches, RL is able to incorporate long-term effects of actions and supports the development of self-learning capabilities in dynamic or partially unknown settings.

Specialized benchmarking libraries and modeling platforms, such as algorithm configuration environments and automated optimization tools, are highly valuable in these tasks. Advanced search methods for multi-dimensional parameter spaces, multi-agent frameworks, and automatic adjustment of agent strategies based on accumulated experience play a decisive role in improving system outcomes. Progress has been made in the development of frameworks for hyperparameter optimization, which can be integrated with RL to further accelerate the search for the most effective configurations.

Modern hardware-software systems present a significant engineering challenge due to their multifaceted structure, the integration of diverse software modules, and their operation in highly variable conditions. A critical issue in their design and evaluation is the effective selection and tuning of input parameters to meet predefined objectives, which often include performance, energy efficiency, and operational stability. Traditional approaches relying on manual configuration or static heuristics frequently fall short in adapting to the dynamic and high-dimensional environments inherent to these systems. Reinforcement learning (RL) offers a transformative framework by enabling agents to learn optimal configuration policies through trial-and-error interaction with the target system, thereby autonomously balancing multiple conflicting objectives. RL agents iteratively adjust parameters based on observed feedback in the form of performance metrics, effectively constructing decision policies that improve over time and are capable of responding to changing conditions and operational uncertainties. The use of dedicated benchmarking libraries and frameworks such as DACBench [1], coupled with modern hyperparameter optimization platforms like Optuna [2] and Auto-WEKA [3], facilitates the systematic evaluation of RL-driven configuration strategies and supports robust and scalable solution development. Moreover, the integration of multi-agent configurations and dynamic parameter control mechanisms enhances the flexibility and scalability of these approaches for complex, large-scale, and distributed environments. This paradigm reduces human intervention, accelerates the search for optimal configurations, and increases the reliability and adaptability of software-hardware systems operating in heterogeneous and evolving contexts.

A reinforcement learning-centered strategy dramatically increases the scalability of testing procedures, ensures robustness to variable external factors, and decreases dependence on subjective human decision making. Such solutions demonstrate substantial potential for adaptation across a wide spectrum of technical problems – from engineering complex information systems to distributed computing platforms. Nevertheless, the successful deployment of RL-based approaches requires in-depth understanding of practical aspects of environment configuration, as well as ongoing evaluation of the performance, stability, and extensibility of known algorithms in various operational contexts.

References:

1. N. Awad, T. Eimer, F. Hutter, M. Lindauer. Multi-agent Dynamic Algorithm Configuration. arXiv preprint. 2022. arXiv:2210.06835.

2. T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama. Optuna: A Next-generation Hyperparameter Optimization Framework. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD'19, Anchorage, AK, USA, 2019. 2623–2631. DOI: 10.1145/3292500.3330701.

3. L. Kotthoff, C. Thornton, H. Hoos, F. Hutter, K. Leyton-Brown. Auto-WEKA 2.0: Automatic model selection and hyperparameter optimization in WEKA. Journal of Machine Learning Research. 2017. 18: 16–261. https://www.jmlr.org/papers/volume18/16-261/16-261.pdf.