Evolutionary Computation and Reinforcement Learning for Cyber-physical System Design

Chengjie Lu
Department of Engineering Complex Software Systems
Simula Research Laboratory
Oslo, Norway
chengjielu@simula.no

***Abstract*—Cyber-physical systems (CPSs) are designed to integrate computation and physical processes through constantly interacting with the physical environment. The complexity and uncertainty of the environment often come up with unpredictable situations, which place high demands on the dynamic adaptability of CPSs. Further, as the environment evolves, the CPS needs to constantly evolve itself to adapt to the changing environment. This paper presents a research plan that aims to develop a novel framework to address CPS design challenges under uncertain environments. We propose to utilize evolutionary computation and reinforcement learning techniques to design control policies that can adapt to the dynamic changes and uncertainties of the environment. Further, novel testing and evaluation approaches that can generate test cases while adapting to dynamic changes in the system and the environment will be explored.**

***Index Terms*—cyber-physical system, evolutionary computation, reinforcement learning, uncertainty**

I. Introduction

Cyber-physical systems (CPSs), which combine computing, networking, and physical processes, have attracted much attention for their ability to provide intelligent control and decision-making in complex and dynamic environments [1]. A typical CPS mainly consists of physical, communication, computation, and control components [2], where the computation component is the core of the whole system and embeds the decision-making strategy. CPSs have demonstrated great potential in various applications, including smart manufacturing, air transportation, critical infrastructure (e.g., smart grid), intelligent transportation and robotics [3]. Such applications often place high requirements on the reliability and robustness of CPSs, which introduces practical challenges for CPS design.

In general, CPS is designed to automatically perform different tasks under dynamic physical environments, which requires the system to be able to adapt to changes in the system and environment states. In addition, the design of an increasingly autonomous CPS is required to consider various situations that may happen in real-world tasks. However, the physical environment is complex and full of various uncertainties, which may bring unpredictable situations; and the increasing complexity and uncertainty of the system often come up with unexpected and unpredictable behaviors of CPSs. Moreover, coordination among multiple agents in a CPS or different CPSs is important for decision-making in complex tasks. This is because coordination enables agents to achieve a common goal and to make decisions based on the actions and decisions of other agents in the system, which is critical for dealing with uncertainties and ensuring the reliability of the system.

Evolutionary computation (EC), inspired by the principles of biological evolution to find the optimal solutions, has been successfully applied in various areas [4]. Applications of EC include design optimization of complex systems [5] by searching for the best system variables and test generation for dynamic systems using multi-objective optimization [6]–[8]. Reinforcement learning (RL) is about agents learning optimal policies by interacting with unknown environments. RL has been widely used in problems that require dynamic adaptation, such as game playing [9], robotics [10], and autonomous systems [11]. In recent years, RL has also demonstrated great potential in testing and evaluating autonomous systems [12], [13]. However, studies have shown that both EC and RL have their own limitations, and several studies have combined these two techniques and proposed evolutionary reinforcement learning (ERL) to address these limitations, while the application of ERL in CPS design has not been extensively studied.

Being aware of the above design challenges of CPSs and the successful applications and limitations of EC and RL, this paper presents a research plan that focuses on utilizing EC and RL to facilitate the design and validation process of CPSs. Concretely, this plan aims at developing a framework that can support the design process of CPSs under uncertain environments. Additionally, we plan to develop new methods to test and evaluate whether a CPS meets design requirements.

II. Research Questions and Expected Contributions

We present the research questions to be addressed and discuss the expected contributions of each research question below.

RQ1: How can we utilize EC and RL to design CPSs that can adapt and evolve while considering the uncertainty and complexity of the environment? This research question aims to address CPS design challenges and improve the ability of CPSs to deal with uncertain and dynamic-evolving environments. The expected contributions include a CPS design framework that can take advantage of EC and RL. Besides, we also consider proposing novel hybrid algorithms that

This work is supported by the Co-evolver project (No. 286898/F20) funded by the Research Council of Norway.

leverage EC and RL, thereby mitigating the limitations of their individual application. As discussed earlier, studies have been conducted and ERL approaches have been proposed to combine the advantages of EC and RL. For example, EC can potentially improve RL's exploration ability, given its ability to optimize in parallel through population evolution. Therefore, this study will investigate the potential application of ERL to address the CPS design challenges.

RQ2: Considering tasks that require multi-agent coordination and interaction, how can we employ EC and RL to coordinate collaboration among multiple agents while considering the dynamic adaption and evolution of the system? This research question intends to develop multi-agent control policies that can coordinate multiple agents as a team for complex tasks. Specifically, this study will contribute to apply existing multi-agent approaches (e.g., multi-agent RL, multi-agent EC) to design cooperative decision-making strategies for CPS. Further, novel multi-agent solutions will be developed that can coordinate different agents toward a common goal or each agent simultaneously toward different goals.

RQ3: How to effectively and efficiently evaluate the ability of CPSs to deal with dynamic and uncertain environments? Given the complexity and uncertainty of CPS's operating environment, existing CPS testing methods may not be able to deal effectively with such environments. Hence, this research question aims to contribute novel adaptive testing strategies that can adapt to dynamic environments and generate test cases for CPSs. In addition, test optimization methods for efficient CPS testing will be investigated.

III. Related Works

In the literature, various approaches have been proposed for addressing the design challenges of CPSs by employing EC [4], [5], [7], [8]. Schranz et al. [14] proposed to evolve controllers for CPSs through EC, thus enabling CPSs to adapt to dynamic environments. He et al. [4] explored the potential application of EC in addressing security design challenges of CPS, such as cyber attack detection [15]. EC is also employed to facilitate the evaluation process of CPSs. For example, Abdessalem et al. [6] proposed to combine evolutionary algorithms with machine learning techniques to efficiently generate critical test scenarios for autonomous driving systems (ADSs).

Various CPSs design solutions using RL have been proposed [10], [11], [16]. Niroui et al. [17] proposed a RL-based control policy for designing rescue robots operating in unpredictable environments. To address the design challenges of ADS under urban driving environments, Chen [18] first designed a novel visual representation approach to encode low-dimensional latent states. Then several RL algorithms are employed to learn driving policies under various complex urban driving scenarios. Multi-agent coordination and interaction are important features of CPS tasks, and multi-agent RL [19] is a promising solution. For example, Zhou et al. [20] employed multi-agent RL to design behavior policies for ADS, where several agents learn to cooperate to ensure the fuel efficiency, comfort, and safety of autonomous driving.

EC and RL have shown successful applications in many challenging CPS design tasks. However, both of these approaches have limitations: EC is typically limited by the complexity of the problems, especially those that require optimization in large parameter spaces [21]. On the contrary, RL can learn optimal behavior policies under complex state spaces, while its effectiveness is limited by sparse rewards and lack of effective exploration [22]. Recently, evolutionary reinforcement learning (ERL) [22]–[24] has been proposed by leveraging the strengths of EC and RL to address their limitations. For example, Khadka and Tumer [22] employed evolutionary algorithms to provide diversified data to train agents in RL, which is effective for sparse rewards problems and can effectively explore the parameter space. ERL methods have been applied to benchmark tasks requiring continuous control, but rarely in the context of CPS design.

IV. Experiment Design and Evaluation

Experiment platform. We plan to conduct the experiment and evaluate the results in the domain of autonomous driving. ADS is a typical CPS capable of sensing the environment and making decisions without human intervention. These systems place high requirements on safety and reliability and usually consist of multiple components to perform different tasks. This study will focus on performing experiments and evaluations in simulated environments. For simulating such an environment, we plan to utilize high-fidelity simulators, such as *Carla* [25]. In addition, we will build a physical experiment platform with three rovers in our laboratory. The rover is equipped with sensors, actuators and computers to support the development of various autonomous driving decision-making strategies.

Experiment design. We plan to develop novel autonomous driving policies and testing strategies, and design experiments accordingly. Concretely, we will employ different driving tasks to evaluate the ability of the driving policy to adapt to various environments and evolve as the environment changes. Further, we will select appropriate multi-agent tasks to evaluate the multi-agent policy, such as ensuring comfort and safety while navigating to the destination. Finally, the developed ADS testing strategy will be evaluated using various ADSs. State-of-the-art approaches will be used for comparisons.

Empirical evaluation. We will select proper metrics to evaluate our approaches. The metrics could include safety or comfort measures of ADSs (e.g., collision, time to collision), and quality measures of the applied algorithms (e.g., accuracy, recall rate [26] and inverted generational distance [27]). Also, we will analyze and report results from different perspectives, using different analytical methods, to draw solid conclusions. In addition, to evaluate whether the proposed approaches are effective in real-world driving, we will study the transferability of virtual-world to physical-world driving. The approaches will be evaluated with three physical rovers in our lab and we will study the ability to perform multi-agent tasks with physical rovers. Finally, we will also study the usage of evaluation approaches in hardware-in-loop testing and physical testing, which is an important industrial need.

References

[1] R. Baheti and H. Gill, "Cyber-physical systems," *The impact of control technology*, vol. 12, no. 1, pp. 161–166, 2011.

[2] H. Chen, "Applications of cyber-physical system: a literature review," *Journal of Industrial Integration and Management*, vol. 2, no. 03, p. 1750012, 2017.

[3] V. Gunes, S. Peter, T. Givargis, and F. Vahid, "A survey on concepts, applications, and challenges in cyber-physical systems," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 8, no. 12, pp. 4242–4268, 2014.

[4] H. He, C. Maple, T. Watson, A. Tiwari, J. Mehnen, Y. Jin, and B. Gabrys, "The security challenges in the iot enabled cyber-physical systems and opportunities for evolutionary computing & other computational intelligence," in *2016 IEEE congress on evolutionary computation (CEC)*, pp. 1015–1021, IEEE, 2016.

[5] A. Arias-Montano, C. A. C. Coello, and E. Mezura-Montes, "Multiobjective evolutionary algorithms in aeronautical and aerospace engineering," *IEEE transactions on evolutionary computation*, vol. 16, no. 5, pp. 662–694, 2012.

[6] R. Ben Abdessalem, S. Nejati, L. C. Briand, and T. Stifter, "Testing advanced driver assistance systems using multi-objective search and neural networks," in *Proceedings of the 31st IEEE/ACM international conference on automated software engineering*, pp. 63–74, 2016.

[7] R. B. Abdessalem, S. Nejati, L. C. Briand, and T. Stifter, "Testing vision-based control systems using learnable evolutionary algorithms," in *Proceedings of the 40th International Conference on Software Engineering*, pp. 1016–1026, 2018.

[8] S. Nejati, "Testing cyber-physical systems via evolutionary algorithms and machine learning," in *2019 IEEE/ACM 12th International Workshop on Search-Based Software Testing (SBST)*, pp. 1–1, IEEE, 2019.

[9] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[10] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, M. Hasan, B. C. Van Essen, A. A. Awwal, and V. K. Asari, "A state-of-the-art survey on deep learning theory and architectures," *electronics*, vol. 8, no. 3, p. 292, 2019.

[11] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. Al Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2021.

[12] C. Lu, Y. Shi, H. Zhang, M. Zhang, T. Wang, T. Yue, and S. Ali, "Learning configurations of operating environment of autonomous vehicles to maximize their collisions," *IEEE Transactions on Software Engineering*, vol. 49, no. 1, pp. 384–402, 2022.

[13] B. Chen, X. Chen, Q. Wu, and L. Li, "Adversarial evaluation of autonomous vehicles in lane-change scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10333–10342, 2021.

[14] M. Schranz, W. Elmenreich, and M. Rappaport, "Designing cyber-physical systems with evolutionary algorithms," *Cyber-physical laboratories in engineering and science education*, pp. 111–135, 2018.

[15] A. Di Giorgio, A. Pietrabissa, F. Delli Priscoli, and A. Isidori, "Robust protection scheme against cyber-physical attacks in power systems," *IET Control Theory & Applications*, vol. 12, no. 13, pp. 1792–1801, 2018.

[16] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238–1274, 2013.

[17] F. Niroui, K. Zhang, Z. Kashino, and G. Nejat, "Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 610–617, 2019.

[18] J. Chen, B. Yuan, and M. Tomizuka, "Model-free deep reinforcement learning for urban autonomous driving," in *2019 IEEE intelligent transportation systems conference (ITSC)*, pp. 2765–2771, IEEE, 2019.

[19] L. Buşoniu, R. Babuška, and B. De Schutter, "Multi-agent reinforcement learning: An overview," *Innovations in multi-agent systems and applications-1*, pp. 183–221, 2010.

[20] W. Zhou, D. Chen, J. Yan, Z. Li, H. Yin, and W. Ge, "Multi-agent reinforcement learning for cooperative lane changing of connected and autonomous vehicles in mixed traffic," *Autonomous Intelligent Systems*, vol. 2, no. 1, p. 5, 2022.

[21] M. J. Kochenderfer, T. A. Wheeler, and K. H. Wray, *Algorithms for decision making*. MIT press, 2022.

[22] S. Khadka and K. Tumer, "Evolutionary reinforcement learning," *arXiv preprint arXiv:1805.07917*, vol. 223, 2018.

[23] S. Khadka, S. Majumdar, T. Nassar, Z. Dwiel, E. Tumer, S. Miret, Y. Liu, and K. Tumer, "Collaborative evolutionary reinforcement learning," in *International conference on machine learning*, pp. 3341–3350, PMLR, 2019.

[24] P. Chrabaszcz, I. Loshchilov, and F. Hutter, "Back to basics: Benchmarking canonical evolution strategies for playing atari," *arXiv preprint arXiv:1802.08842*, 2018.

[25] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on robot learning*, pp. 1–16, PMLR, 2017.

[26] T. Vafeiadis, K. I. Diamantaras, G. Sarigiannidis, and K. C. Chatzisavvas, "A comparison of machine learning techniques for customer churn prediction," *Simulation Modelling Practice and Theory*, vol. 55, pp. 1–9, 2015.

[27] S. Ali, P. Arcaini, D. Pradhan, S. A. Safdar, and T. Yue, "Quality indicators in search-based software engineering: an empirical evaluation," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 29, no. 2, pp. 1–29, 2020.