Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper, eds.

REINFORCEMENT LEARNING WITH DISCRETE EVENT SIMULATION: THE PREMISE, REALITY, AND PROMISE

Sahil Belsare
Mohammad Dehghanimohammadabadi

Mechanical and Industrial Engineering Department
Northeastern University
360 Huntington Avenue
Boston, MA 02115, USA

ABSTRACT

Recently, Reinforcement Learning (RL) has been successfully applied in domains like manufacturing, supply chain, health care, finance, robotics, and autonomous vehicles. For such applications, uncertainty in the real-life environment presents a significant challenge in training an RL agent. One of the approaches to tackle this obstacle is by augmenting RL with a Discrete Event Simulation (DES) model. This framework enables us to accommodate the list of possible circumstances occurring in the practical environment. In this presentation, we analyze the existing literature of RL models using DES to put forward the benefits, avenues, challenges, and scope for future work in developing such models for industrial optimization.

1 NEED OF SIMULATION FOR REINFORCEMENT LEARNING

Traditional optimal control methods, although being efficient, account for limited online computation capabilities, prove ineffective for large scale problems due to the curse of dimensionality and the curse of modeling. RL's ability of 'self-learning algorithms' for sequential decision making has become an attractive solution to overcome these setbacks. This is evident by the growing body of literature that is applying RL with DES technique to existing Industrial and Operations Research problems, as seen in figure 1. Furthermore, for complex industrial optimization, training of the RL agent may require hundreds of thousands of steps to achieve even a near-optimal policy, thus making it infeasible to train the agent on physical setup. The other challenge is to explore an environment, where the transition probabilities are unknow (Feldkamp et al. 2020). Such complex models can be augmented with RL in a simulated environment for training an agent. Following on that, the papers were selected based on the inclusion criteria, that is to have usage of RL with DES for industrial application. The papers were obtained from the database of Winter Simulation Conference (WSC), Elsevier, SIMULTECH, IEEE, and arXiv. A total of 50 papers were selected for the abstract review stage and 23 were selected for complete reading.

2 RL WITH DES FOR INDUSTRIAL OPTIMIZATION

The majority of the research has been published on applications of RL with DES in the domain of manufacturing, followed by the supply-chain industry. This popularity arises since both industries are highly dependent on the policy in effect (Arulkumaran et al. 2017). Studies have shown successful results in training an agent to achieve an average lower cost of action and effectively identify optimal operating policies of production facilities. An adaptive order release mechanism based on Deep Q Reinforcement Learning with a simulation environment outperformed the benchmark set by the traditional method by yielding lower total costs, less mean and standard deviation of tardiness, and a shorter shop floor throughput time. An important characteristic of a simulated environment is that it randomly samples new states from the simulation model. This enables the RL+DES model to map the stochasticity of the real-life environment, a highly important characteristic for Job-shop problems . Solving Job-shop problems with Deep Q-Network (DQN) and DES have shown significant savings in the overall system's cost and are able to find an optimal policy of dispatching rules once trained using the integrated environment. The next popular topic of interest is policy development for autonomous vehicles in factories. It is in such cases where simulation facilitates an agent to interact with the environment in a digital world and gain its intelligence while saving time and money (Agrell et al. 2021). Researchers can model real life constraints like narrow warehouse aisles, multiple pick up and drop off locations, the possibility of collision, and restricted traffic routing. When benchmarked with the shortest travel distance rule, the DQN model reduced the occurrences of traffic congestion. To add, obtaining real world data to train an RL agent for supply chain optimization can be inconvenient considering the time spent to acquire such a dataset by performing an experiment (Rabe and Dross. 2015). Researchers have shown improvement in inventory management when optimized using DES and RL framework for supply chain problems.

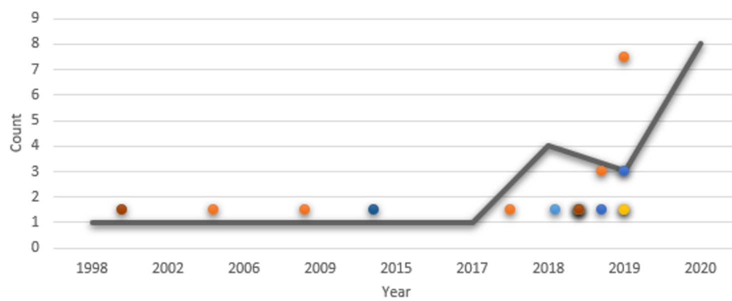

Figure 1: Increase in publications on RL with DES for Industrial Optimization.

3 CONCLUSION

A compelling case to combine simulation with RL to overcome the uncertainties in the industrial application was made by the reviewed studies. A significant challenge in expanding research on RL with DES is that the practitioners in each area have backgrounds in using different tools/software and thus, there is a requirement to learn skill sets from both areas. It is followed by the challenge of developing a standardized way to evaluate and benchmark RL techniques against traditional known methods. There is limited literature available on reviewing the combined use of RL with DES models, leaving a large gap for further research. Thus, the noticeable success in implementing an integrated framework calls for a need to analyze its vast potential. The parallel goal of this presentation is to encourage researchers to work on the challenges addressed. This needs to be accomplished by a combined effort from the simulation, reinforcement learning, and operations research community.

REFERENCES

Agrell, C., K. R. Dahl, and A. Hafver. 2021. "Optimal sequential decision making with probabilistic digital twins". arXiv preprint arXiv:2103.07405.

Arulkumaran, K., M. P. Deisenroth, M. Brundage, and A. A. Bharath. 2017. "Deep Reinforcement Learning: A Brief Survey". IEEE Signal Processing Magazine 34(6):26–38.

Feldkamp, N., S. Bergmann, and S. Strassburger. 2020. "Simulation-Based Deep Reinforcement Learning For Modular Production Systems". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T.Roeder, and R.Thiesing, eds, 1596– 1607. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rabe, M., and F. Dross. 2015. "A reinforcement learning approach for a decision support system for logistics networks". In Proceedings of the 2015 Winter Simulation Conference, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds, 2020–2032. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.