Integrating Reinforcement Learning and Discrete Event Simulation for Enhanced Production Scheduling

Arthur Santâna da Silva[1][0000-0001-9867-9426]; Pedro Amaral Pereira [2][0009-0005-44464-120X]; Eduardo Félix de Simas Mauger Canova[1][0000-0003-3958-5555]; Manoel Carlos Pego Saisse[2][0009-0000-7352-927X]; Andréa Regina Nunes de Carvalho[2][0000-0002-8294-4552].

[1] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[2] National Institute of Technology, Rio de Janeiro, Brazil

Abstract. Production scheduling involves planning activities and allocating resources to meet customer demands effectively. Real production scheduling problems tend to be very complex due to the large number of variables, constraints and objectives (which can occasionally conflict with one another) involved. Automated approaches, like simulation-based systems utilizing Discrete Event Simulation (DES), can be useful in reducing the weight of these complexities. DES may mimic real-world production systems as logically separated processes occurring at specific times, subjected to a set of constraints that can be adjusted according to the environment being simulated. But, even with the assistance of DES, defining the set of operational decisions that will lead to the best production schedule, according to the objectives of the organization, remains a challenge, especially considering the explosive combinatorial nature of production scheduling problems. To address this challenge, Reinforcement learning (RL) has been combined with DES, notably in the context of Industry 4.0. Since, in RL, the learning process resembles the DES approach when the human agent goes through a process of trial and error to guarantee the highest possible reward (the best result from a combination of performance indicators), an expected result is the extraction of knowledge about which actions must be taken (which dispatching rules must be chosen to which machines) when a certain scenario (a state in RL) is identified.

This study refers to an experimental project that integrates a finite capacity DES-based technology with RL to support production scheduling. Differently from most strategies proposed in the literature, where the RL algorithms are trained to define each task allocation to resources individually, throughout the execution of the DES, the agent evaluates alternative production schedules built by the DES based on different operational decisions (such as dispatching rules) in search of the best set that will lead to the enhancement of the pursued objectives. Starting from a reference DES system, the project was organized into three phases: RL modeling, implementation of the DES-RL based model, and case study experimentations.

The DES system considered in this study equip schedulers with the capability to build alternative schedules by simulating the application of operational decisions such as prioritizing production orders, adjusting capacity (e.g., through overtime, subcontracting), leveraging flexibility (e.g., reallocating tasks to alternative resources), and adapting material availability and promised delivery dates (e.g., renegotiating delivery dates with suppliers or customers). Diverse performance indicators such as service level, lead-time, operational costs, and resource utilization are calculated at the end of each simulation run. To emulate this relationship between the human scheduler and the DES system, at each loop, the agent takes one or more actions in the form of operational decisions. After the simulation run, a reward is calculated and awarded to the agent. Presently, the range of actions available to the agent is limited to prioritizing tasks. This involves determining which task will be processed when multiple tasks compete for the same resources at the same moment of the planning horizon. The value

of the reward provided to the agent is based on the overall profit calculated by de DES. The state representation was modeled as a vector with n positions, where the first n/2 positions refer to prioritizing production orders and the second half deprioritizing these orders. Each position denoted a particular task to be scheduled. A toy case study was conducted to test the DES-RL based model involving a job shop environment with 5 machines and 5 production orders with different routings, comprising 24 activities, all competing for the same 5 machines. Monetary penalties for delays are incurred, and the objective is to maximize the profit, considering given operational costs. Numeric results shows that DQN agent, randomly, managed to obtain the optimal reward value for the problem before 2,000 iterations on average (less than 1 minute). However, it still struggles to converge to this result, even after 100,000 iterations with a few different combinations of hyperparameters, as the limited action and state space may hinder the agent from learning and improving a policy. Considering this limiting factor, the model still needs to be optimized by revisiting the structure of action, the state modeling and the scope of algorithm hyperparameters.

For future studies, the set of actions allowed to be taken by the agent could be broadened, considering the operational decisions cited above, already included in the DES system. It is crucial to identify the right characteristics of the production environment that must be included in the DES, avoiding extremes (overrepresentation and underrepresentation) that could hinder the agent learning process. The reward function could include various performance measures already calculated by the DES system (e.g., service level, resource utilization, profit). Other frameworks of RL such as Graph Neural Networks can also be applied to exploit the disjunctive graph representation of production scheduling problems. This study showcases the potential of RL together with DES as a viable approach for developing production scheduling tools, enhancing operational efficiency in manufacturing industries. For future steps, the aim is to explore more complex case studies with additional production specificities and the full range of decisions available in the DES system.

Keywords: Production Planning, Reinforcement Learning, Discrete Event Simulation.

References

1. A. N. Carvalho, L. F. Scavarda, and L. J. Lustosa, “Implementing finite capacity production scheduling: lessons from a practical case”, International Journal of Production Research, vol. 52, no. 4, pp. 1215–1230, 2014.
2. A. Esteso, D. Peidro, J. Mula, and M. Díaz-Madroñero, “Reinforcement learning applied to production planning and control,” International Journal of Production Research, vol. 61, pp. 1–18, 08 2022.
3. M. Parente, G. Figueira, P. Amorim, and A. Marques, “Production scheduling in the context of industry 4.0: review and trends,” International Journal of Production Research, vol. 58, no. 17, pp. 5401–5431, 2020.
4. C. Zhang, W. Song, Z. Cao, J. Zhang, P. S. Tan, and X. Chi, “Learning to Dispatch for Job Shop Scheduling via Deep Reinforcement Learning” In Advances in Neural Information Processing Systems, vol. 33, pp. 1621–1632, 2020.