OFFICIAL TEMPLATE FOR THE SUBMISSION OF PAPERS
II INTERNATIONAL SCIENTIFIC CONVENTION
"II ICC UCLV 2019"

JUNE 23th – 30th, 2019
CAYOS DE VILLA CLARA. CUBA.

II Conferencia Internacional de Procesamiento de la Información (CIPI 2019). International Workshop on Internet of Things and Artificial Intelligence (IOTAI 2019)

Multi-Agent Reinforcement Learning Tool for Job Shop Scheduling Problems

Jessica Coto Palacio[1], Yailen Martínez Jiménez[2], Ann Nowé[3]

1-Jessica Coto Palacio. UEB Hotel Los Caneyes, Cuba, informatico@caneyes.vcl.tur.cu
2-Yailen Martínez Jiménez. Universidad Central "Marta Abreu" de Las Villas, Cuba, yailenm@uclv.edu.cu
3- Ann Nowé. Vrije Universiteit Brussel, Bégica, ann.nowe@ai.vub.ac.be

Abstract: The emergence of Industry 4.0 allows for new approaches to solve industrial problems such as the Job Shop Scheduling Problem. It has been demonstrated that Multi-Agent Reinforcement Learning approaches are highly promising to handle complex scheduling scenarios. In this paper we propose a user friendly Multi-Agent Reinforcement Learning tool, more appealing for industry. It allows the users to interact with the learning algorithms in such a way that all the constraints in the production floor are carefully included and the objectives can be adapted to real world scenarios. The user can either keep the best schedule obtained by a Q-Learning algorithm or adjust it by fixing some operations in order to meet certain constraints, then the tool will optimize the modified solution respecting the user preferences using two possible alternatives. These alternatives are validated using OR-Library benchmarks, the experiments show that the modified Q-Learning algorithm is able to obtain the best results.

Keywords: Job Shop Scheduling; Industry 4.0; Reinforcement Learning, Multi-Agent Systems.

During the last years, technological developments have increasingly benefited industry performance. The appearance of new information technologies has given rise to intelligent factories in what is termed as Industry 4.0 or i4.0 (Leitao et al. 2016; Leitao et al. 2005). The i4.0 revolution involves the combination of intelligent and adaptive systems using shared knowledge among diverse heterogeneous platforms for computational decision-making within Cyber-Physical Systems (CPS) (Leusin et al. 2018; Vogel-Heuser et al. 2015). In this sense, embedding Multi-Agent Systems into CPS is a highly promising approach to handle complex and dynamic problems. A typical example of an industrial opportunity of this kind is scheduling, whose goal is to achieve resource optimization and minimization of the total execution time (Toader 2017).

Given the complexity and dynamism of industrial environments, the resolution of this type of problem may involve the use of very complex solutions, as customer orders have to be executed, and each order is composed by a number of operations that have to be processed on the resources or machines available. In real world scheduling problems, the environment is so dynamic that all this information is usually not known beforehand. For example, manufacturing scheduling is subject to constant uncertainty, machines breakdown, orders take longer than expected, and these unexpected events can make the original schedule fail (Hall and Potts 2004; Xiang and Lee 2008). Accordingly, the problem of creating a job-shop scheduling, known as Job-Shop Scheduling Problem (JSSP), is considered one of the hardest manufacturing problems in literature (Asadzadeh 2015).

Different Operations Research (OR) techniques (Linear Programming, Mixed-Integer Programming, etc) have been applied to scheduling problems. These approaches usually involve the definition of a model, which contains an objective function, a set of variables and a set of constraints. OR based techniques have demonstrated the ability to obtain optimal solutions for well-defined problems, but OR solutions are restricted to static models. Artificial Intelligence approaches, on the other hand, provide more flexible representations of real-world problems, allowing human expertise to be present in the loop

(Gomes 2000; Martínez Jiménez 2012a; Zhang 1996). In (Gabel and Riedmiller 2007; Gabel 2009), the authors suggested and analyzed the application of Reinforcement Learning (RL) techniques to solve job shop scheduling problems. They demonstrated that interpreting and solving this kind of scenarios as a multi-agent learning problem is beneficial for obtaining near-optimal solutions and can very well compete with alternative approaches. These agents typically use RL, which is learning what to do (how to map situations to actions) so as to maximize a numerical reward signal (Sutton and Barto 1998). It allows an agent to learn optimal behavior through trial-and-error interactions with its environment. By repeatedly trying actions in different situations the agent can discover the consequences of its behavior and identify the best action for each situation. For example, when dealing with unexpected events, learning methods can play an important role as they could 'learn' from previous results and change specific parameters for next iterations, allowing not only to find good solutions but more robust ones.

Another problem that has been identified in the scheduling community is the fact that most of the research concentrates on optimization problems that are a simplified version of reality. As the author points out in (Urlings 2010): "this allows for the use of sophisticated approaches and guarantees in many cases that optimal solutions are obtained. However, the exclusion of real-world restrictions harms the applicability of those methods. What the industry needs are systems for optimized production scheduling that adjust exactly to the conditions in the production plant and that generate good solutions in very little time". In this research we propose a Multi-Agent Reinforcement Learning (MARL) tool that allows the user to either keep the best result obtained by a learning algorithm or to include extra constraints of the production floor. This first version allows to fix operations to time intervals in the corresponding resources and afterwards optimize the solution based on the new constraints added by the user. This is a first approach that helps to close the gap between literature and practice.

The MARL tool groups several algorithms aimed at solving scheduling problems in the manufacturing industry. It focuses on the need of building a more flexible schedule, in

order to adjust it to the user's requests without violating the restrictions of the JSSP scenario. The approach used to obtain the original solution that the user can afterwards modify is the one proposed in (Martínez Jiménez 2012), it is a generic multi-agent reinforcement learning approach that can easily be adapted to different scheduling settings, such as the Flexible Job Shop (FJSSP) or the Parallel Machines Job Shop Scheduling (PMJSSP). The algorithm used is the Q-Learning, which works by learning an action-value function that gives the expected utility of taking a given action in a given state. There is basically an agent per machine which takes care of allocating the operations that must be executed by the corresponding resource.

Once the user chooses the scheduling problem to solve, the tool proposes an initial solution based on the original QL algorithm mentioned before, and at the same time it enables a set of options that are the basis of this research. The user has the possibility to move the operations either using the mouse or the touch screen, and these movements must be validated once the new positions are decided.

In this work we compare the performance of two alternatives for optimizing the schedule once the user has fixed some operations, the classical left shifting and a modified Q-Learning algorithm, which includes the position of the fixed operations in the learning process. In order to measure the performance of the alternatives several benchmark problems from the OR-Library (Beasley 1990) were used. For each instance the same operations were fixed, and each optimization alternative had to adjust the schedule in order to minimize the makespan. The Wilcoxon test applied to the results shows that there are significant differences between the two alternatives (sig=0.08), the mean ranks confirm that the QL version with fixed operations is able to obtain better results than the classical optimization process. This is mainly because the left shifting respects the order in which the operations were initially placed along the x axis. The QL algorithm, on the other hand, keeps the fixed positions and during the process of learning, the order in which the operations are scheduled in the resources does not have to be the same, this allows the approach to obtain better solutions in terms of makespan.

References

Asadzadeh, Leila. 2015. “A Local Search Genetic Algorithm for the Job Shop Scheduling Problem with Intelligent Agents.” *Computers & Industrial Engineering* 85: 376–383.

Beasley, J. E. 1990. “OR-Library: Distributing Test Problems by Electronic Mail.” *Journal of the Operational Research Society* 41(11): 1069–1072.

Gabel, Thomas. 2009. “Multi-Agent Reinforcement Learning Approaches for Distributed Job-Shop Scheduling Problems.” PhD Thesis, Universität Osnabrück.

Gabel, Thomas, and Riedmiller, M. 2007. “Scaling Adaptive Agent-Based Reactive Job-Shop Scheduling to Large-Scale Problems.” In *IEEE Symposium on Computational Intelligence in Scheduling (CI-Sched 2007)*, 259–266.

Gomes, Carla P. 2000. “Artificial Intelligence and Operations Research: Challenges and Opportunities in Planning and Scheduling.” *The Knowledge Engineering Review* 15(1): 1–10.

Hall, N G, and Potts, C. N. 2004. “Rescheduling for New Orders.” *Operations Research* 52: 440–453.

Leitao, P et al. 2005. “Intelligent Products: The Grace Experience.” *Control Engineering Practice* 42: 95–105.

Leitao, P, Colombo, A. W., and Karnouskos, S. 2016. “Industrial Automation Based on Cyber-Physical Systems Technologies: Prototype Implementations and Challenges.” *Computers Industry* 81: 11–25.

Leusin, Matheus E et al. 2018. “Solving the Job-Shop Scheduling Problem in the Industry 4.0 Era.” *Technologies* 6(4). http://www.mdpi.com/2227-7080/6/4/107.

Martínez Jiménez, Yailen. 2012. “A Generic Multi-Agent Reinforcement Learning Approach for Scheduling Problems." PhD Thesis, Vrije Universiteit Brussel.

Sutton, Richard S., and Barto, Andrew G. 1998. *Reinforcement Learning: An Introduction*. The MIT Press.

Toader, Florentina Alina. 2017. “Production Scheduling in Flexible Manufacturing Systems: A State of the Art Survey.” 3(7): 1–6.

Urlings, Thijs. 2010. “Heuristics and Metaheuristics for Heavily Constrained Hybrid Flowshop Problems.” Universidad Politécnica de Valencia.

Vogel-Heuser, B, J Lee, and P Leitao. 2015. “Agents Enabling Cyber-Physical Production Systems.” *AT-Autom.* 63: 777–789.

Xiang, W, and H. P. Lee. 2008. “Ant Colony Intelligence in Multi-Agent Dynamic Manufacturing Scheduling.” *Engineering Applications of Artificial Intelligence* 21: 73–85.

Zhang, Wei. 1996. “Reinforcement Learning for Job Shop Scheduling.” PhD Thesis. Oregon State University.