A Summary Of Using Reinforcement Learning Strategies For Treating Project And Production Management Problems

Georgios Koulinas
Department of Production and Management Engineering
Democritus University of Thrace
Xanthi, Greece
gkoulina@pme.duth.gr

Alexandros Xanthopoulos
Department of Production and Management Engineering
Democritus University of Thrace
Xanthi, Greece
axanthop@pme.duth.gr

Athanasios Kiatipis
Fujitsu Technology Solutions GmbH
Munich, Germany
athanasios.kiatipis@ts.fujitsu.com

Dimitrios Koulouriotis
Department of Production and Management Engineering
Democritus University of Thrace
Xanthi, Greece
jimk@pme.duth.gr

***Abstract*— Recently, Reinforcement Learning (RL) strategies have attracted researchers' interest as a powerful approach for effective treating important problems in the field of production and project management. Generally, RL are autonomous machine learning algorithms that include a learning process that interacts with the problem, which is under study in order to search for good quality solutions in reasonable time. At each decision point of the algorithm, the current state of the problem is revised and decisions about the future of the searching strategy are taken. The objective of this work is to summarize, in brief, recently proposed studies using reinforcement learning strategies for solving project scheduling problems and production scheduling problems, as well. Based on the review, we suggest directions for future research about approaches that can be proved interesting in practice.**

Keywords— Reinforcement Learning, Project Scheduling, Production Scheduling, Resource Constrained Project Scheduling Problem, Job Shop Scheduling Problem

I. Introduction

The present paper summarizes the preliminary study of the scientific literature relative to the usage of reinforcement learning strategies for treating classic project management and production management problems, as well. More specifically, we provide a sample of papers that have applied RL to project and production management problems, rather than generating an exhaustive survey. This study aims to investigate the recently proposed trends in this field.

The rationale for this research paper is the following. Firstly, the fields of project and production scheduling are closely linked in terms of optimization problem formulations, solution approaches and so forth. Over the past few decades, a significant number of papers that address project/production scheduling problems with RL-based methods have been published. Nonetheless, a relevant survey that would categorize and analyze existing works is absent from the literature. This is in contrast to other application domains of RL such as robotics, traffic and railways scheduling etc.

In order to find the relative recent studies, we used the keywords "reinforcement learning", "project management", and "production management", in the title, abstract or keywords. Also, we have included only journal papers and conference proceedings written in English, and published between the years 2004 to 2018 for project management and between the years 1997 to 2018 for production management field. Initially, there were 94 papers initially retrieved and finally, due to their relevancy 33 of them, included. The search has been realized using the popular online database, Scopus. The choices and the search string for extracting the relative papers are summarized in the following Table 1.

In this study we have organized the reviewed papers according to the general problem category that is under study. Thus, we grouped papers that use RL approaches to treat project scheduling problems such as the Resource Constrained Project Scheduling Problem (RCPSP) and some of its variations. Also, we collected papers that employ a reinforcement learning process to solve a variety of production planning problems and their extensions such as the Job-Shop Scheduling Problem, the Flow-Shop Scheduling Problem.

TABLE I. REVIEW CHOICES AND SEARCH STRING

Item	Choice	
Keywords	Project management reinforcement learning, project scheduling	Production management reinforcement learning, production scheduling
Boolean Operators	AND between keywords	AND between keywords
Search fields	Title; Abstract; Keywords	Title; Abstract; Keywords
Exclusion criteria	Papers that do not address the main topic	Papers that do not address the main topic
Language	English	English

978-1-5386-5244-2/18/$31.00 ©2018 IEEE

Publication type	Journal or Conference proceedings	Journal or Conference proceedings
Time window	2004 – 2018	1997 - 2018
Search string	TITLE-ABS-KEY (reinforcement AND learning AND project AND scheduling) AND (LIMIT-TO (SRCTYPE , "p") OR LIMIT-TO (SRCTYPE , "j"))	TITLE-ABS-KEY (reinforcement AND learning AND production AND scheduling) AND (LIMIT-TO (SRCTYPE , "p") OR LIMIT-TO (SRCTYPE , "j"))

II. PROJECT SCHEDULING

The Resource Constrained Project Scheduling Problem (RCPSP) is a classic project management problem that has attracted the rising interest of researchers and practitioners. Historically, it is one of the most important and challenging resource and task allocation problems. The reviewed studies that developed algorithms used reinforcement learning for treating this class of problem are discussed in the present section. Generally, there were 11 papers included, studying RL applications to project management problems. The following Fig. 1 shows their allocation over the previous years. As illustrated, the plenty of papers have been published during the last 7 years proving that applying RL to project management problems has not attracted much of research interest yet.

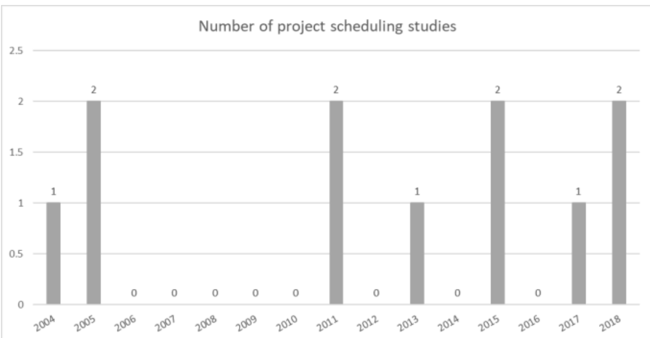

Fig. 1. Number of papers treating project management problems

In [1] presented a general methodology to improve the results of the Regret-Based Biased Random Sampling Scheme (RBRS). More specifically, they mixed the rout-algorithm with the support vector machine (SVM) to learn a value function which directs the search strategy given by RBRS. The proposed approach allowed decreasing the size of the training set and showed better performance.

The study of [2] used the rout-algorithm, to adapt simple greedy heuristics for the RCPSP. In addition, several methods for defining the value function are used and the support vector machine (SVM) for approximation purposes.

The reference [3] focused on using simulation techniques for creating scheduling policies, in order to save computation time. Also, techniques based in reinforcement learning were used to empower the simulation possibility in the field of software cybernetics.

In [4] presented a multi-agent reinforcement learning approach for constructing competitive solutions for the multi-mode resource-constrained project scheduling problem (MRCPSP). In this study, a network of distributed reinforcement learning agents, each one assigned to an activity, in order to learn to select successor activities. The results of applying this approach proved to be competitive against the best performing algorithms for the MRCPSP.

The study of [5] employed reinforcement learning for defining optimal schedules for software development projects. This approach includes the construction of a stochastic scheduling model, simulation of the project, and task assignments analysis. Among their most important findings was that optimal scheduling policies assign tasks according to previous performance of the resources and the specific characteristics of the software.

In the reference [6] proposed an approach to enhance the functionality of a metaheuristic algorithm by using reinforcement learning to generate heuristics.

The reference [7] proposed an approach including a team of asynchronous agents (A-Team), implemented using a multi-agent system, for solving the Multi-mode Resource-Constrained Project Scheduling Problem (MRCPSP). The A-Team consists of multiple agents and shared memory which during interactions, managed by a dynamic learning strategy, produce different solutions to the problem. The proposed reinforcement learning based dynamic learning strategy, operates on interactions among agents and the shared memory.

In [8] suggested an approach that involves a group of project managers, each one learning an activity list using reinforcement learning, and especially learning automata, whilst they are learning to select an appropriate place to put each activity list among the other activity lists. The results of this study have shown that applying this scheme clearly improves the objective of minimizing the project delay.

In the study of [9] proposed a teaching–learning-based optimization algorithm (TLBO) to treat the multi-skill resource constrained project scheduling problem (MS-RCPSP) with the makespan minimization criterion. In this study, a reinforcement phase is included into the TLBO with a permutation and a resource-based local search as improvement process to the local intensification.

In the reference [10] analyzed the application of a Monte Carlo Tree Search (MCTS) supported by the Upper Confidence Bounds Applied to Trees (UCT) method, (MCTS/UCT), combined with deep neural networks trained with a reinforcement learning algorithm, to real-world problems such as the Capacitated Vehicle Routing Problem and the Risk-Aware Project Scheduling Problem.

The study of [11] applied a team of asynchronous agents method and population based methods to treat the multi-mode resource-constrained project scheduling problem. As agents served algorithms like local search, path relinking, and tabu search, that cooperate through reinforcement learning, and population learning. The results are evaluated using nonparametric statistical tests.

The following Table 2 summarizes the reviewed papers according to the algorithmic approach used, and the project management problem that is under study.

TABLE II. PAPERS TREATING PROJECT MANAGEMENT PROBLEMS

Problem	Reference	Algorithmic method
RCPSP	[1]	Rout algorithm of RL with SVM
RCPSP	[2]	Rout algorithm of RL with SVM
Scheduling software engineering activities	[3]	RL & simulation
MRCPSP	[4]	Multi-agent RL strategies
Scheduling software development activities	[5]	RL & simulation
Decentralized resource-constrained multi-project scheduling problem	[6]	RL & metaheuristic algorithm
MRCPSP	[7]	A-team multi-agent system
MRCPSP	[8]	Learning automata
Multi-skill RCPSP	[9]	TLBO
CVRP & risk RCPSP	[10]	MCTS/UCT, deep neural networks & RL
MRCPSP	[11]	Asynchronous agents and population methods

III. PRODUCTION SCHEDULING

Reinforcement learning approaches have been widely used in the past for addressing production scheduling problems. Typically, such type of problems consists of jobs that are assigned to resources at specific times. The objective is to minimize the total processing time, namely the makespan. The proposed algorithms for this type of problems are discussed in the following section. In brief, 22 papers surveyed, presenting RL applications to production management problems. In Fig. 2 illustrated the number of papers appeared each year. As showed, RL in production scheduling is used during the last 20 years but the last decade there is a rising research interest in this topic.

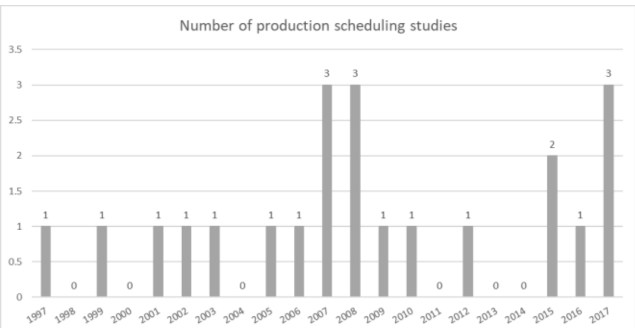

Fig. 2. Number of papers treating production scheduling problems

The work of [12] designed and applied a hybrid neural network for a rescheduling problem in production. The proposed approach has been tested against other methods for treating production rescheduling problems.

In the study of [13] proposed a neural network that optimizes dispatching rules according to an optimization goal. This approach was tested on two benchmark scheduling problems.

The reference [14] presented a distributed learning and control (DLC) approach for a multi-objective scheduling problem. The experimental results show that the proposed algorithm is more efficient than simple dispatching rules in just-in-time production systems problems.

In [15] proposed an approach using reinforcement learning agent and a simulation model. This approach has been tested on a classic scheduling problem.

In the study of [16] used reinforcement Learning enhanced by simulation to solve a multi-period scheduling problem in a two-stage production process.

The reference [17] proposed composite rules, for machine and job selection, trying to threat a dynamic job-shop scheduling problem. A reinforcement learning (RL) algorithm used to enhance its adaptive and learning capability. The results with the RL-based approach illustrates that it can be employed as a real-time scheduler.

In [18] presented an algorithm that operates in a market-based production control system. This approach consists of a simulated annealing algorithm, a reinforcement learning system, and an artificial neural network.

In the study of [19] used a reinforcement learning method (R-learning) to treat parallel machine scheduling problems for minimizing mean flow time of jobs. Also, a heuristic algorithm proposed that uses two very popular dispatching rules such as the Shortest Expected Processing Time (SEPT) and First In First Out (FIFO) rule. The main result was that the reinforcement learning method was proved to be more efficient than heuristic rules in all test problems.

In [20] applied Q-learning to a dynamic single-machine scheduling problem. The Q-learning employed to select an appropriate dispatching rule dynamically. The experimental results showed that, Q-learning is a promising approach for treating the agent-based dynamic production scheduling problem.

In the study of [21] presented a learning approach for approximating the failure rate of a production system and then trying to find the trade-off that maximizes system's potentially good results. The efficiency of this study was tested in an automated deadlock-detection system.

In the study of [22] proposed a model for job-shop scheduling considering uncertainty, and a hybrid algorithm combining Simulation, Neural Networks and Genetic Algorithm. The efficiency of the approach is checked through a case example.

The study of [23] used reinforcement learning to find out proper scheduling rules for selecting intermediate factories in production networks framework. Numerous experiments performed, in order to confirm whether proper rules can be learned.

In the study of [24] presented an evolutionary algorithm for the non-stationary bandit problem. The performance of the proposed method were validated using simulation-based experiments.

The reference [25] proposed a dynamic adaptive iterative scheduling for the dynamic job-shop scheduling problem. The authors tried to transform the scheduling problem into reinforcement learning problems, easier to treat. The experimental results of the analysis were promising.

The study of [26] introduced the use of average-reward reinforcement learning for simulation-based development and adaptation of a manner for dynamic order acceptance under uncertainty in make-to-order manufacturing systems. The feasibility of the method illustrated through comparisons with different order acceptance heuristics.

In [27] proposed a relational learning process to update rescheduling knowledge that can be used in reactive scheduling. This approach was tested in an industrial example with very good results.

In the study of [28] presented a Reinforcement Learning algorithm (Q-Learning) to solve Job Shop and Flow Shop scheduling problems. The application results were validated using test cases previously published and compared with the reported optimal values.

The reference [29] proposed an approach for optimally scheduling a multi-stage and multi-machine production system for multiple types of products. Experimental results showed that the proposed method provided more reliable than the other heuristic and multi-agent algorithms.

In [30] proposed an RL process to assist the cooperative function of the staff,together with machines. The system keeps the schedules informed according to changing real-time events.

The reference [31] proposed a heterarchical approach based on a reinforcement Q-learning approach, and applied the most suitable Machine Selection Rule and Dispatching Rule to deal with complex scheduling problems.

In [32] used reinforcement learning to develop a gantry scheduling policy in order to improve system production, given that the gantry has to move materials among machines and buffers. The gantry, using Q-Learning algorithm, learns how find the best movement and generally how to act under various situations in order to minimize system production loss. For validating the efficiency of the proposed algorithm, a comparison with "First Come First Serve" policy has been made.

Finally, [33] introduced a reinforcement learning-based approach for minimizing two contradictory objectives: the average inventory level and the average number of backorders. The efficiency of the proposed process investigated through simulation experiments.

The following Table 3 summarizes the reviewed papers according to the algorithmic approach used, and the production management problem that is under study.

TABLE III. PAPERS TREATING PRODUCTION MANAGEMENT PROBLEMS

Problem	Reference	Algorithmic method
Rescheduling of batch manufacture problem	[12]	Hybrid neural network
Job-shop scheduling (minimize total tardiness)	[13]	Neural network based agent
Real-time shop floor control problem	[14]	Distributed learning and control (DLC)
Batch sizing and job sequencing in unreliable, multiproduct serial line	[15]	Reinforcement learning and simulation
Production planning in a two-stage manufacturing process	[16]	Reinforcement learning and Monte Carlo simulation
Dynamic job-shop scheduling problem	[17]	Q-learning and simulation
Job scheduling in a market-based production control system	[18]	Simulated annealing algorithm, a reinforcement learning system, and an artificial neural network
Parallel machine scheduling problems	[19]	R-learning, simulation and a heuristic method
Dynamic single-machine scheduling problem	[20]	Q-learning algorithm
Failure detection scheduling in an online production system	[21]	ALG algorithm and simulation
Job-shop scheduling problem	[22]	Simulation, Neural Networks and Genetic Algorithm
Production networks scheduling	[23]	Reinforcement Learning (Profit sharing) approach
Non-stationary bandit problem	[24]	Evolutionary algorithm
Dynamic job-shop scheduling problem	[25]	Dynamic adaptive iterative scheduling
Order admission control in make-to-order manufacturing system	[26]	Average-reward reinforcement learning
Reactive production planning	[27]	Relational learning process
Job Shop and Flow Shop	[28]	Q-Learning algorithm

scheduling problems		
Multi-stage and multi-machine manufacturing system scheduling	[29]	Single-agent AQL (approximate Q-learning), multi-agent Q-learning and simulation
Multi-stage and multi-machine manufacturing system scheduling with multi-skilled workforce	[30]	Multi-agent reinforcement learning approach
Partially flexible job-shop scheduling	[31]	Q-learning algorithm
Gantry scheduling problem	[32]	Q-Learning algorithm
Minimizing the average inventory level and the average number of backorders	[33]	R-Learning algorithm

IV. CONCLUSIONS

In this paper, 33 articles about the use of RL techniques for treating project management and production management problems are reviewed. RL has been used widely for production management but not as much for project scheduling for which, the majority of papers have been published during the last seven years. As a result, this can be an interesting future research topic, since there is much experience of applying such learning strategies to production scheduling. Also, the authors believe that it would be proved quiet interesting to further develop neural networks approaches for various project scheduling problems, and also, hybrid nature based metaheuristics enhanced by a reinforcement learning process. In terms of learning algorithms it is seen that the standard Q-learning method is adopted in numerous publications. Given the plethora of discounted/average reward and model-based/model-free algorithms, it would be interesting to examine alternative, state-of-the-art approaches. Furthermore, it is noted that the major advantage of RL over traditional methods such as Dynamic Programming is that it overcomes the "dimensionality curse" and/or the lack of analytical models of the controlled systems. Consequently, it is interesting to apply RL methods in large-scale systems where analytical models are not available.

ACKNOWLEDGMENT

The work of A. Kiatipis was supported by the BigStorage: Storage-Based Convergence Between HPC and Cloud to Handle Big Data project from the European Union through the Marie Skłodowska-Curie Actions framework under Grant H2020-MSCA-ITN-2014-642963.

REFERENCES

[1] K. Gersmann and B. Hammer, "A reinforcement learning algorithm to improve scheduling search heuristics with the SVM," in *IEEE International Conference on Neural Networks - Conference Proceedings*, 2004, vol. 3, pp. 1811–1816.

[2] K. Gersmann and B. Hammer, "Improving iterative repair strategies for scheduling with the SVM," *Neurocomputing*, vol. 63, no. SPEC. ISS., pp. 271–292, 2005.

[3] F. Padberg, "On the potential of process simulation in software project schedule optimization," in *Proceedings - International Computer Software and Applications Conference*, 2005, vol. 2, pp. 127–130.

[4] T. Wauters, K. Verbeeck, G. V. Berghe, and P. De Causmaecker, "Learning agents for the multi-mode project scheduling problem," *J. Oper. Res. Soc.*, vol. 62, no. 2, pp. 281–290, 2011.

[5] F. Padberg and D. Weiss, "Optimal scheduling of software projects using reinforcement learning," in *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*, 2011, pp. 9–16.

[6] T. Wauters, K. Verbeeck, P. De Causmaecker, and G. Vanden Berghe, *Boosting metaheuristic search using reinforcement learning*, vol. 434. 2013.

[7] P. Jędrzejowicz and E. Ratajczak-Ropel, *Reinforcement learning strategy for solving the MRCPSP by a team of agents*, vol. 39. 2015.

[8] T. Wauters, K. Verbeeck, P. De Causmaecker, and G. Vanden Berghe, "A learning-based optimization approach to multi-project scheduling," *J. Sched.*, vol. 18, no. 1, pp. 61–74, Feb. 2015.

[9] H.-Y. Zheng, L. Wang, and X.-L. Zheng, "Teaching–learning-based optimization algorithm for multi-skill resource constrained project scheduling problem," *Soft Comput.*, vol. 21, no. 6, pp. 1537–1548, 2017.

[10] J. Mańdziuk, *MCTS/UCT in solving real-life problems*, vol. 738. 2018.

[11] E. Ratajczak-Ropel, "Experimental Evaluation of Agent-based Approaches to Solving Multi-mode Resource-Constrained Project Scheduling Problem," *Cybern. Syst.*, pp. 1–21, 2018.

[12] M. I. Heywood, M.-C. Chan, and C. R. Chatwin, "Application of stochastic real-valued reinforcement neural networks to batch production rescheduling," *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, vol. 211, no. 8, pp. 591–603, 1997.

[13] S. Riedmiller and M. Riedmiller, "A neural reinforcement learning approach to learn local dispatching policies in production scheduling," in *IJCAI International Joint Conference on Artificial Intelligence*, 1999, vol. 2, pp. 764–769.

[14] J. Hong and V. Prabhu, *Distributed learning and control for manufacturing systems scheduling*, vol. 2070. 2001.

[15] D. C. Creighton and S. Nahavandi, "The application of a reinforcement learning agent to a multi-product manufacturing facility," in *Proceedings of the IEEE International Conference on Industrial Technology*, 2002, vol. 2, pp. 1229–1234.

[16] H. Cao, H. Xi, and S. F. Smith, "A reinforcement learning approach to production planning in the fabrication/fulfillment manufacturing process," in *Winter Simulation Conference Proceedings*, 2003,

vol. 2, pp. 1417–1423.
[17] Y.-Z. Wei and M.-Y. Zhao, "Reinforcement learning-based approach to dynamic job-shop scheduling," *Zidonghua Xuebao/Acta Autom. Sin.*, vol. 31, no. 5, pp. 765–771, 2005.
[18] B. C. Csáji, L. Monostori, and B. Kádár, "Reinforcement learning in a distributed market-based production control system," *Adv. Eng. Informatics*, vol. 20, no. 3, pp. 279–288, 2006.
[19] Z.-C. Zhang, L. Zheng, and M. X. Weng, "Parallel machines scheduling with reinforcement learning," *Jisuanji Jicheng Zhizao Xitong/Computer Integr. Manuf. Syst. CIMS*, vol. 13, no. 1, pp. 110–116, 2007.
[20] S.-J. Wang, S. Sun, B.-H. Zhou, and L.-F. Xi, "Q-learning based dynamic single machine scheduling," *Shanghai Jiaotong Daxue Xuebao/Journal Shanghai Jiaotong Univ.*, vol. 41, no. 8, 2007.
[21] F. Zeng, "A reinforcement-learning approach to failure-detection scheduling," in *Proceedings - International Conference on Quality Software*, 2007, pp. 161–170.
[22] G. Zhang, C. Li, J. Zhu, and H. Zhu, *Hybrid intelligent algorithm for job-shop scheduling under uncertainty*, vol. 5315 LNAI, no. PART 2. 2008.
[23] H. Yamaba, S. Matsumoto, and S. Tomita, "An attempt to obtain scheduling rules of Network-Based Support System for Decentralized Scheduling of Distributed Production Systems," in *IEEE International Conference on Industrial Informatics (INDIN)*, 2008, pp. 506–511.
[24] D. E. Koulouriotis and A. Xanthopoulos, "Reinforcement learning and evolutionary algorithms for non-stationary multi-armed bandit problems," *Appl. Math. Comput.*, vol. 196, no. 2, pp. 913–922, 2008.
[25] Y. Wei, X. Jiang, P. Hao, and K. Gu, "Pattern driven dynamic scheduling approach using reinforcement learning," in *Proceedings of the 2009 IEEE International Conference on Automation and Logistics, ICAL 2009*, 2009, pp. 514–519.
[26] F. Arredondo and E. Martinez, "Learning and adaptation of a policy for dynamic order acceptance in make-to-order manufacturing," *Comput. Ind. Eng.*, vol. 58, no. 1, pp. 70–83, 2010.
[27] J. Palombarini and E. Martínez, "SmartGantt - An interactive system for generating and updating rescheduling knowledge using relational abstractions," *Comput. Chem. Eng.*, vol. 47, pp. 202–216, 2012.
[28] Y. C. Fonseca Reyna, Y. Martínez Jiménez, J. M. Bermúdez Cabrera, and B. M. Méndez Hernández, "A reinforcement learning approach for scheduling problems," *Investig. Operacional*, vol. 36, no. 3, pp. 225–231, 2015.
[29] S. Qu, T. Chu, J. Wang, J. Leckie, and W. Jian, "A centralized reinforcement learning approach for proactive scheduling in manufacturing," in *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, 2015, vol. 2015–Octob.
[30] S. Qu, J. Wang, S. Govil, and J. O. Leckie, "Optimized Adaptive Scheduling of a Manufacturing Process System with Multi-skill Workforce and Multiple Machine Types: An Ontology-based, Multi-agent Reinforcement Learning Approach," in *Procedia CIRP*, 2016, vol. 57, pp. 55–60.
[31] W. Bouazza, Y. Sallez, and B. Beldjilali, "A distributed approach solving partially flexible job-shop scheduling problem with a Q-learning effect," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 15890–15895, 2017.
[32] J. Arinez, X. Ou, and Q. Chang, "Gantry scheduling for two-machine one-buffer composite work cell by reinforcement learning," in *ASME 2017 12th International Manufacturing Science and Engineering Conference, MSEC 2017 collocated with the JSME/ASME 2017 6th International Conference on Materials and Processing*, 2017, vol. 4.
[33] A. S. Xanthopoulos, A. Kiatipis, D. E. Koulouriotis, and S. Stieger, "Reinforcement Learning-Based and Parametric Production-Maintenance Control Policies for a Deteriorating Manufacturing System," *IEEE Access*, vol. 6, pp. 576–588, 2017.