Available online at www.sciencedirect.com
ScienceDirect
Procedia CIRP 99 (2021) 27–32
www.elsevier.com/locate/procedia

14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP ICME '20

# Foresighted digital twin for situational agent selection in production control

Marvin Carl May[a,*], Leonard Overbeck[a], Marco Wurster[a], Andreas Kuhnle[a], Gisela Lanza[a]

[a] *wbk Institute of Production Science, Karlsruhe Institute of Technology (KIT), Kaiserstr. 12, 76131 Karlsruhe, Germany*

* Corresponding author. Tel.: +49-1523-950-2624; *E-mail address:* marvin.may@kit.edu

**Abstract**

As intelligent Data Acquisition and Analysis in Manufacturing nears its apex, a new era of Digital Twins is dawning. Foresighted Digital Twins enable short- to medium-term system behavior predictions to infer optimal production operation strategies. Creating up-to-the-minute Digital Twins requires both the availability of real-time data and its incorporation and serve as a stepping-stone into developing unprecedented forms of production control. Consequently, we regard a new concept of Digital Twins that includes foresight, thereby enabling situational selection of production control agents. One critical element for adequate system predictions is human behavior as it is neither rule-based nor deterministic, which we therefore model applying Reinforcement Learning. Owing to these ever-changing circumstances, rigid operation strategies crucially restrain reactions, as opposed to circumstantial control strategies that hence can outperform traditional approaches. Building on enhanced foresights we show the superiority of this approach and present strategies for improved situational agent selection.


Peer-review under responsibility of the scientific committee of the 14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 15-17 July 2020.

*Keywords:* Fluid Automation; Production Control; Digital Twin; Machine Learning; Human Behavior

## 1. Motivation

In the wake of increased data availability a surge in the application of Industry 4.0 applications can be observed [1]. A major driver for such applications is the rising demand for individualized products at industrialized costs achieved through "volume-related economies" [2]. Hence, production system flexibility and agility is known as a key requirement, inducing the need for powerful production control methods [3]. While traditional production control by means of advanced heuristics or mathematical optimization is easily comprehensible, modern learning techniques, such as reinforcement learning based [4] or monte carlo tree search based [5], can improve overall production system performance. However, solely using these control forms comes at the expense of further disadvantages, as indicated in Table 1. Thus, the individualized evaluation of production control methods is necessary in order to find the most suitable technique for the production system and circumstances at hand.

The idea of Cyber-Physical Production Systems (CPPS) has long aimed at providing transparency, allowing real-time production control, empowering the Internet of Things (IoT) and production data analyses [6, 7]. Digital Twins that extend this notion become increasingly popular fields of research [8], as they enable to check production system conformance to product requirements. Further advantages are the simplification and speeding up of data acquisition, production system planning [6] and acting as an enabler towards the digital transformation [9] which resulted in high prospects [7]. As opposed to the application of Digital Twins for future predictions for production planning [6], the introduction of foresight for near-real-time decisions is largely unexplored. Thus, this paper addresses this research gap by presenting a framework for production control agent selection by means of situational digital twin foresight.

2212-8271 
Peer-review under responsibility of the scientific committee of the 14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 15-17 July 2020.
10.1016/j.procir.2021.03.005

Table 1. Comparison between production control approaches for Industry 4.0

| Criterion | Mathematical Optimization | Heuristics | Reinforce-ment Learning | Monte Carlo Tree Search |
|---|---|---|---|---|
| Single Objective | ● | ● | ● | ● |
| Global Perspective | ● | ◐ | ● | ● |
| Flexibility | ○ | ◐ | ◐ | ◐ |
| Adaptability | ○ | ◐ | ◐ | ● |
| Real-time Operation | ○ | ● | ● | ◐ |
| Computing Power | high | low | medium | medium |
| New Situation Handling | ◐ | ◐ | ● | ◐ |
| Robustness | ○ | ◐ | ● | ◐ |
| Customization | ● | ◐ | ● | ● |
| Reliability | ● | ◐ | ◐ | ◐ |
| Comprehensibility | ◐ | ● | ○ | ◐ |

Legend: ○:not fulfilled - ◐:partially fulfilled - ●:fulfilled

The remainder of this paper is structured as follows: Within Section 2 Production Planning and Control (PPC) as well as Digital Twins are defined, followed by Section 3 which presents the methods for advancing digital twins with foresight. The case study and its results are presented in Section 4. The paper is concluded with a discussion and outlook in Section 5.

## 2. Digitalized Production Planning and Control

Production Planning and Control (PPC) serves a holistic approach incorporating time, financial flow and material flow for the optimization of production systems [10]. Traditional PPC has been focused around the three partially segregated dimensions: (1) strategic, (2) tactical and (3) operational decision making time-horizon [11]. However, the subdivision into three sub-problems followed by separate optimizations may not be well suited for the adaptability required by agile and changeable production systems. The integration of technological solutions to tear down the separating walls of traditional PPC can mitigate the influences of time-wise partitioning and pave the way towards a truly Smart Factory [12]. In particular, on an operational level production control can integrate both the analytical capabilities of Digital Shadows and the real-time capabilities of production system Digital Twins [12].

### *2.1. Digital Shadow*

A Digital Shadow in production system context "provides a holistic concept for a manufacturing-oriented information (supply) system" [13] and additionally serves as an interface connecting data storage, acquisition and warehouse [14] for production oriented data analytics. Despite technical challenges and differences in concrete implementations, the main concept is to feasibly link the information generated

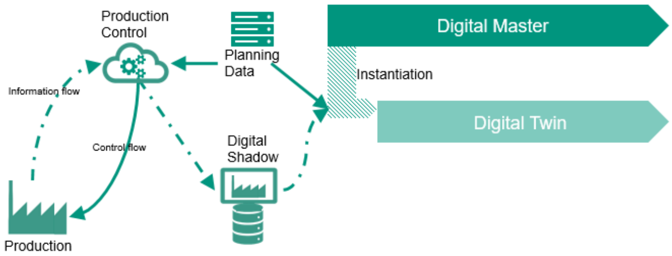


Fig. 1. Instantiation of digital twins for production systems

within the production system as well as the information control and feedback [13]. In particular the introduction of automated data acquisition, data connection and correlation [14] enables the analysis of physical processes in a cyber sphere. In doing so, the system's status at any point within the entire history can be digitally accessed [12]. Thus, creating up-to-the-minute digital shadows enables the latent and up-to-date description of production systems states.

### *2.2. Digital Twin*

While the definition of a Digital Twin is incomplete, limited [8] and extremely diverse [9], herein the following definition is used: "A Digital Twin is" the digital instantiation "of a unique (physical) asset" with "similar properties, conditions and behavior" [8]. The exact technical realization however may differ [15], yet the underlying structure consisting of a Digital Master and a Digital Shadow [8] remains. The Digital Master can launch several instances of Digital Twins based on the information released from planning data and the corresponding timely information given by the Digital Shadow as shown in Figure 1. Creating Digital Twins in hindsight allows an in-depth analysis of past events through data analytics. Thus, optimization through simulations can be delegated to far more accurate Digital Twins, which reflect the reality in great detail and accuracy. Digital Twins, thus, are successors of common manually designed simulation models, which have a high degree of abstraction and, therefore, restricted validity [6]. In creating a more realistic environment than regular simulations can provide, the identification of more suitable production control methods flourishes. Nonetheless, the fact that an up-to-the minute Digital Twin can provide accurate insight about the production system's short- and medium-term behavior is yet to be exploited.

## 3. Foresighted Digital Twin

This section presents the conceptual inclusion of foresight into Digital Twins and outlines the effective modeling of human behavior within production systems by means of reinforcement learning.

### *3.1. Including foresight for situational control agent selection*

Key in evaluating a digital twin is its ability to correctly depict any situation taking place in the underlying physical

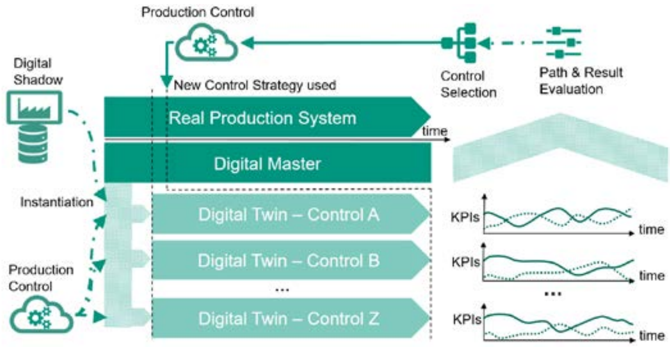


Fig. 2. Foresighted digital twins evaluation for situational production system control selection

system [7]. Not until a point in time is portrayed can the dynamics of the production system be included. Instead of following regular Digital Twin simulation goals of predicting long-term system behavior accurately [6], in other words the dynamic's outcome, foresighted Digital Twins allow a glimpse at the concrete path the dynamics most likely follow. Hence, alleviating the gap in short- and medium-term predictions. However, current approaches to Digital Twins which include foresight are limited to big data predictions of a production systems status [16] and machine process predictions [17].

Owing to the ability of changeable and agile production systems to reconfigure, to select a fluid level of automation and produce inherently different products, individual production targets can easily differ among different occasions. Yet, rigid production control policies and time-invariant strategies are incapable of dealing with such frequent changes. In this case any decision making instance within a Digital Twin is modeled as an agent, in particular production control. Hence, the circumstantial near real-time control agent selection in foresighted digital twins can alleviate the effects of such non target-aligned production control. In a similar vein, sudden and unexpected changes in the production system can induce the necessity to adapt to more suitable production control policies.

Including foresight and path evaluation itself is by no means a novel concept as it is at the heart of Monte Carlo methods. Instead of following random roll-outs to determine system behavior [18, 19] or creating an asymmetric, diverging tree as in MCTS [20], the foresighted Digital Twin enables a more targeted control policy search. Hence, given the ability of a foresighted Digital Twin to accurately reflect the influence of different control strategies on the systems development, their near real-time comparison becomes meaningful for accurate system predictions and agent selection as shown in Figure 2. In a similar vein to Monte Carlo methods, the core of situational control agent selection in Digital Twins is the instantiation of several Digital Twins based on up-to-the minute information each with a different control strategies running in parallel. In contrast to current Digital Twin developments, however, the evaluation is based on the current situation, the systems development path and final status. As a result the control strategy most suitable to current circumstances can be selected.

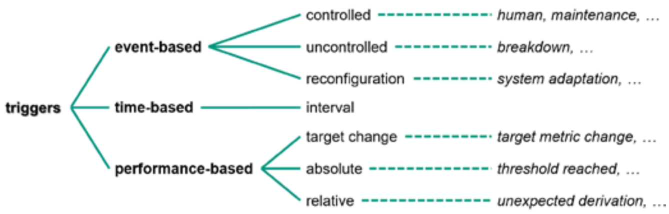


Fig. 3. Categorization of triggers starting foresight instantiation

Moreover, the concept of triggering foresight and selecting new control policies can be separated into subcategories. Most importantly, the instantiation of foresighted Digital Twins can be activated by events, time-based triggers or performance evaluation triggers as visualized in Figure 3. The former can be disaggregated into external events, such as changes in product mix, internal events, such as failures or maintenance, reaching particular situations or defined status of the system as well as metrics defined on changes between two consecutive system statuses. Performance-based triggers can set off the comparison of control policies in order to sustain certain performance levels or react to a change in production targets. Moreover, the evaluation and selection of control policies can be based on path, result, system and human evaluations. However, in this paper not the full extent of control policy switching modalities and foresighted Digital Twins is explored.

### *3.2. Modeling Human Behavior through Reinforcement Learning*

The influence of human behavior and human resource management strategies on manufacturing performance is widely accepted [21]. In particular high control and commitment strategies increase workers performance [21] which is increased with the ability of individuals to simulate others' decisions [22]. While the latter has been studied using reinforcement learning [22, 23], human behavior in general can effectively be modeled through Markov models [24] or in a more modern approach with reinforcement learning [25]. Despite the inability of reinforcement learning to perfectly match the human learning process [25] its application provides an alternative to system dynamics modeling of human behavior [26] which can be used for goal-oriented human control tasks. Its ability to successfully control production systems [27] and maintenance [28] has been shown.

In accordance to [29] $\langle S, A, P, r\rangle$ represents a standard Markov Decision Process (MDP) where $S$ represents the state space, $A$ the action space, $P$ the stochastic transition function $p(s'|s,a)$ and $r(s,a)$ the reward function defining the reward for action $a$ in state $s$ [29]. Based on individual goals, such as the drive to finish early, deliver high quality and take breaks, the reward function can be individualized to accurately reflect the behavior of individuals. An individualized policy $\pi$ is used to select an action according to $a_t \sim \pi(\cdot\,|s_t)$, depending on experience tuples$(s_t, a_t, r_t, s_{t+1})$.

## 4. Case study: exemplary job-shop

This section presents the case study for the application of agent selection in foresighted Digital Twins based on a simplified exemplary job-shop.

### *4.1. Human Behavior job-shop use-case*

The regarded use-case consists of a matrix shaped layout with 2 × 3 production cells. This matrix production system is controlled by three Automated Guided Vehicles (AGV) and a human supervisor who occasionally engages in transport operations, as shown in Figure 4.

Human behavior is approximated through a Trust Region Policy Optimization (TRPO) reinforcement learner [30] and a weighted reward function including covered distance and system performance. The exact weight is adjusted to correspond to human behavior manually, yet with meaningful and sufficient data an automated approach can be realized through learning from demonstration. Within this simplified control setting for Digital Twins with foresight, the following exemplary order dispatching production control methods are implemented: First-In-First-Out Heuristics (FIFO), Shortest Queue heuristic (SQ), Nearest-Job-First (NJF) a parametrized Composite Rule including distance, queue length and first-in with equal weights (CR) and a reinforcement learning policy (RL). This use-case uses a combination of weighted Key Performance Indicators (KPI) value and timely pareto-comparison of control strategies for final control strategy selection which is performed through a human production manager.

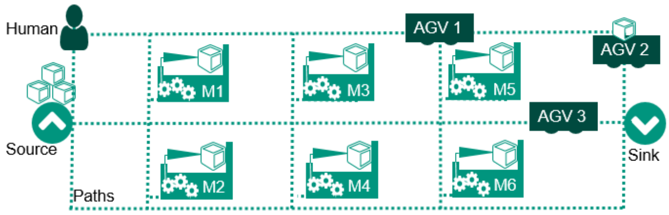


Fig. 4. The regarded job-shop based matrix production system use-case

### *4.2. Analysis and Results*

Based on the performance realized by each policy within the foresighted Digital Twins, their detriments and merits can be assessed. In this case the comparison is based on the systems throughput, average waiting time per finished order and AGV utilization. The individual performance over time and averages are reported exemplarily in Figure 5: A sudden failure of M2 in the given exemplary job-shop triggers the comparison by means of foresighted Digital Twins at time 0 leading to increasing waiting times throughout the period.

In contrast to findings for comparable job-shops and control through different heuristics and reinforcement learners as in [4, 27], the FIFO heuristic does not minimize the average waiting time in this time-frame. In order to control this production system meaningfully within the foresight period one would choose SQ realizing high throughput and short waiting times. In addition, preference-based selection can take the entire run of the KPI curves into account, i.e. consider CR for improved throughput and only few waiting time derivations. Thus, the superiority of situational agent selection instead of long-term

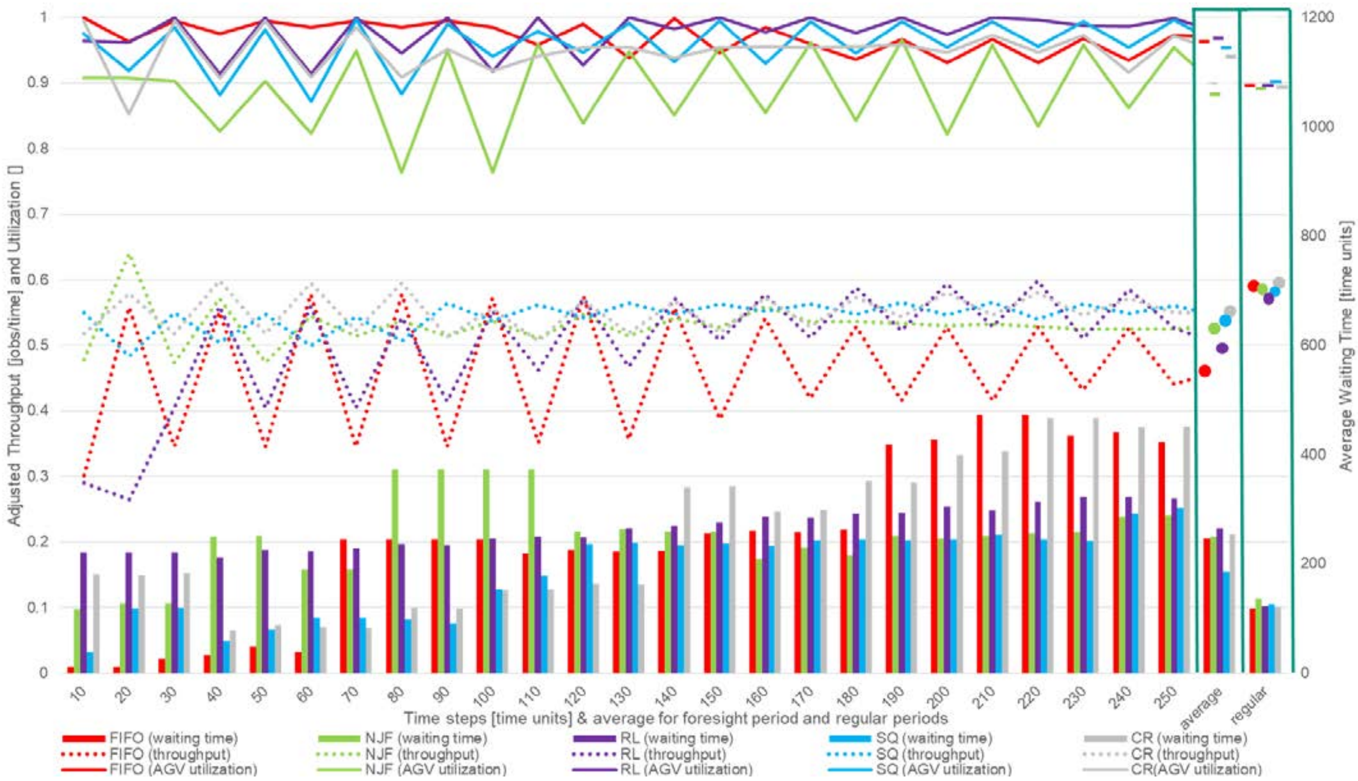


Fig. 5. Comparison of different production control methods and their performances in foresighted digital twins

optimization of production systems can be shown, as the finally selected control strategy (SQ) is dominated by other control strategies for regular operation (see Figure 5 bottom right). The quest in production control, hence, shifts from finding the best control strategy towards selecting the best control strategy for each situation. No control policy is continuously dominating all others in the sense that it is performing better at each point in time. Therefore, a situational control adaption to the currently best suited policy will outperform any pure policy. In other words, according to the "no-free-lunch theorem" no universal optimization strategy outperforms strategies that specialize in particular problems [31] and circumstantial agent selection chooses the most suitable at a given time.

In general, throughout the analysis the high influence of current circumstances on different performance KPIs becomes apparent. As learning techniques are hardly capable of dealing with completely unknown situations, production control selection in foresighted Digital Twins can bridge the time gap arising while learning optimized control policies. Despite the ability to generalize the tremendous changeability of smart manufacturing production systems makes exploring all possible states and generalization infeasibly complex.

## 5. Discussion and Outlook

In a nutshell, this paper introduces and applies a novel concept, foresighted Digital Twins, in order to assess the influence of different production control policies on system performance and behavior in near real-time and select the most suitable strategy for the situation at hand. By modeling human behavior through a reinforcement learner the unpredictability of human behavior is anticipated. The application to a job-shop shows the ability of foresighted Digital Twins to select suitable control strategies. However, the inclusion of final state evaluation remains to be analyzed and can increase the concept's applicability. In order to improve the benefits of foresighted Digital Twins the inclusion of further policies and their validation in a real-world application is advisable.

Most promising for further research is the in-depth analysis of different foresight triggers and the comprehensive evaluation of control policy switching modalities. The latter as well as a breakdown of circumstantial features and selected policies can generate new insights about the regarded production systems. Further research could apply foresighted Digital Twins to different near real-time control problems.

## Acknowledgements

This research work was undertaken in the context of the DIGIMAN4.0 project ("DIGItal MANufacturing Technologies for Zero-defect Industry 4.0 Production", http://www.digiman4-0.mek.dtu.dk/). DIGIMAN4.0 is a European Training Network supported by Horizon 2020, the EU Framework Programme for Research and Innovation (Project ID: 814225).

## References

[1] Lasi, H., Fettke, P., Kemper, H.-G., Feld, T. *et al.*, 2014. Industry 4.0. Business & information systems engineering *6*, p. 239.

[2] Duray, R., Ward, P.T., Milligan, G.W., Berry, W.L., 2000. Approaches to mass customization: configurations and empirical validation. Journal of operations management *18*, p. 605.

[3] Zawadzki, P., Żywicki, K., 2016. Smart product design and production control for effective mass customization in the Industry 4.0 concept. Management and Production Engineering Review *7*.

[4] Stricker, N., Kuhnle, A., Sturm, R., Friess, S., 2018. Reinforcement learning for adaptive order dispatching in the semiconductor industry. CIRP Annals *67*, p. 511.

[5] Edelkamp, S., Gath, M., Greulich, C., Humann, M. *et al.*, 2016. Monte-Carlo tree search for logistics, in *Commercial Transport*, Springer, p. 427.

[6] Uhlemann, T.H.-J., Lehmann, C., Steinhilper, R., 2017. The digital twin: Realizing the cyber-physical production system for industry 4.0. Procedia CIRP *61*, p. 335.

[7] Negri, E., Fumagalli, L., Macchi, M., 2017. A review of the roles of digital twin in cps-based production systems. Procedia Manufacturing *11*, p. 939.

[8] Stark, R., Kind, S., Neumeyer, S., 2017. Innovations in digital modelling for next generation manufacturing system design. CIRP Annals *66*, p. 169.

[9] Kritzinger, W., Karner, M., Traar, G., Henjes, J. *et al.*, 2018. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine *51*, p. 1016.

[10] Wiendahl, H.-P., ElMaraghy, H.A., Nyhuis, P., Zäh, M.F. *et al.*, 2007. Changeable manufacturing-classification, design and operation. CIRP Annals *56*, p. 783.

[11] Wiendahl, H.-P., 1997. *Fertigungsregelung: Logistische Beherrschung von Fertigungsabläufen auf Basis des Trichtermodells*. Hanser München.

[12] May, M.C., Kuhnle, A., Lanza, G., 2020. Digital production and intelligent production control [Digitale Produktion und intelligente Steuerung]. Wt Werkstatttechnik Online *110*, p. 655.

[13] Bauernhansl, T., Hartleif, S., Felix, T., 2018. The Digital Shadow of production - A concept for the effective and efficient information supply in dynamic industrial environments. Procedia CIRP *72*, p. 69.

[14] Schuh, G., Walendzik, P., Luckert, M., Birkmeier, M. *et al.*, 2016. Keine Industrie 4.0 ohne den Digitalen Schatten. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb *111*, p. 745.

[15] Ding, K., Chan, F.T.S., Zhang, X., Zhou, G. *et al.*, 2019. Defining a digital twin-based cyber-physical production system for autonomous manufacturing in smart shop floors. International Journal of Production Research *57*,

p. 6315.
[16] Zhuang, C., Liu, J., Xiong, H., 2018. Digital twin-based smart production management and control framework for the complex product assembly shop-floor. The International Journal of Advanced Manufacturing Technology *96*, p. 1149.
[17] Liu, J., Zhou, H., Liu, X., Tian, G. *et al.*, 2019. Dynamic evaluation method of machining process planning based on digital twin. IEEE Access *7*, p. 19312.
[18] Runarsson, T.P., Schoenauer, M., Sebag, M., 2012. Pilot, rollout and monte carlo tree search methods for job shop scheduling, in *International Conference on Learning and Intelligent Optimization*, p. 160.
[19] Lucas, S.M., Samothrakis, S., Perez, D., 2014. Fast evolutionary adaptation for monte carlo tree search, in *European Conference on the Applications of Evolutionary Computation*, p. 349.
[20] Browne, C.B., Powley, E., Whitehouse, D., Lucas, S.M. *et al.*, 2012. A survey of monte carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in games *4*, p. 1.
[21] Arthur, J.B., 1994. Effects of human resource systems on manufacturing performance and turnover. Academy of Management journal *37*, p. 670.
[22] Suzuki, S., Harasawa, N., Ueno, K., Gardner, J.L. *et al.*, 2012. Learning to simulate others' decisions. Neuron *74*, p. 1125.
[23] Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A. *et al.*, 2009. Reinforcement learning signal predicts social conformity. Neuron *61*, p. 140.
[24] Pentland, A., Liu, A., 1999. Modeling and prediction of human behavior. Neural computation *11*, p. 229.
[25] Shteingart, H., Loewenstein, Y., 2014. Reinforcement learning and human behavior. Current Opinion in Neurobiology *25*, p. 93.
[26] Block, J., Pickl, S., 2014. The mystery of job performance: a system dynamics model of human behavior, in *Proceedings of the 32nd international conference of the System Dynamics Society, Delft, Netherlands, July*, p. 20.
[27] Kuhnle, A., Schäfer, L., Stricker, N., Lanza, G., 2019. Design, Implementation and Evaluation of Reinforcement Learning for an Adaptive Order Dispatching in Job Shop Manufacturing Systems. Procedia CIRP *81*, p. 234.
[28] Kuhnle, A., Jakubik, J., Lanza, G., 2019. Reinforcement learning for opportunistic maintenance optimization. Production Engineering *13*, p. 33.
[29] Sutton, R.S., Barto, A.G., 2018. *Reinforcement learning: An introduction*. MIT Press.
[30] Schulman, J., Levine, S., Abbeel, P., Jordan, M. *et al.*, 2015. Trust region policy optimization, in *International conference on machine learning*, p. 1889.
[31] Ho, Y.-C., Pepyne, D.L., 2002. Simple explanation of the no-free-lunch theorem and its implications. Journal of optimization theory and applications *115*, p. 549.