Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

INTEGRATING SCHEDULING OF LOGISTIC SUPPORT PROCESSES IN AGENT-BASED INDUSTRY 4.0 ASSEMBLY SIMULATION

Adrian Freiter

Department of Logistics
Fraunhofer Institute for
Software and Systems Engineering (ISST)
Speicherstraße 6
44147 Dortmund, GERMANY

Christian Schwede

Faculty of Engineering and Mathematics
University of Applied Sciences and Arts
Bielefeld
Interaktion 1
33619 Bielefeld, GERMANY

ABSTRACT

The upcoming decentralized production systems seem to be promising in Industry 4.0 assembly to handle the challenges of highly individual products. Matrix production characterized by freely linked workstations and an advanced automation level are highly flexible. That is why many efforts have already been made to explore the advantages compared to existing flow shop production systems, but also the additional challenges arising from this new paradigm. One of these challenges is the synchronization of main product and supply part flow at the individual workstations during order scheduling. This paper presents a new approach of integrating logistics support processes into the scheduling of the main product flow to consider the part supply in the decisions taken during scheduling avoiding waiting times. We compare our integrated approach with the existing decoupled scheduling approach, based on a “bicycle assembly” scenario. The results are promising particularly when part supply is a bottleneck.

1 INTRODUCTION

The increasing number of product variants, fluctuations in demand (Kern et al. 2015) in combination with product cycles that are being shortened (Lempp und Siegfried 2022) challenges the automotive and other industry sectors. As a result, manufacturers have adapted their production systems from mass production towards mass customization (Kern et al. 2015). Against this background, the suitability of the traditional production line in the final assembly as the production system of the future is being challenged (Greschke 2016). Matrix production is considered a promising alternative (Göppert et al. 2018; Hofmann et al. 2019) and is based on independent/ freely interlinked modular workstations that are connected to each other by autonomous guided vehicles (AGV) (Peter Greschke 2016). In literature, multi-agent systems (MAS) are preferred as decentral control systems for matrix production in particular due to their scalability (Feldkamp et al. 2019; Mayer et al. 2019). Here, an order is managed by an agent and routed based on the assembly priority graph through the production system. This leads to ad hoc part demands at the workstations. Since single assembly steps of multi-variant products can often only be performed when the individual parts are also available at the workstation. Achieving synchronization between the main product flow (e.g. of the car chassis) and the supporting production logistics flow of the individual parts to avoid waiting times and bad capacity utilization of the workstations is one of the main challenges of controlling these systems. Therefore, logistics, respectively the part supply to the workstations, should be integrated in the scheduling executed by the MAS (Skubowius et al. 2019). Until now to our knowledge there are only two validated approaches considering logistics in a multi-agent decentral control system (Filz et al. 2019; Schmidtke et al. 2023). However, still logistics and production are considered as two separate systems decoupled by a transport order queue leading to poor capacity utilization of resources especially in cases of disruptions when the parts are not available, or transport resources are a limiting factor. Hence closer integration between production and logistics is required. The approach presented in this paper is based on the concept of integrating the part supply in the scheduling, including especially the occupancy and the position of the supply transport means (Blesing et al. 2017). We developed an event-discrete multi-agent simulation environment in Python on top of a software for digital twin-based process planning and control (RIOTANA) from Fraunhofer ISST and use it to compare the decoupled and the integrated scheduling approach based on a scenario of an assembly system for individualized bicycles. We will show that the integrated approach outperforms the decoupled solution especially in cases of limited transportation resources.

The rest of the paper is structured as follows. In section 2 we will give a detailed problem definition of the order scheduling for multi-variant products in a matrix production. In section 3 we will give an overview of the literature and demonstrate the research gap, introducing the decouple control approach in detail. In section 4 we will describe our integrated approach for a control system. In section 5 we will introduce the bicycle assembly system as application case and describe the setup and parameters of the experiments. In section 6 we present the experiments and the results. Section 7 sums up the findings and presents an outlook on further research.

2 PROBLEM DESCRIPTION: ORDER SCHEDULING FOR MULTI-VARIANT PRODUCTS IN A MATRIX PRODUCTION

Multi-variant products have a huge number of configurable variants that are created often based on an assembly of far less variant supply parts, in such a way that from a customer point of view the resulting product seems individual. Products are defined by attributes that are organized in clusters (e.g. a special car engine in the engine cluster). Attributes combined by propositional logic from different clusters are used to define the production steps needed as well as the part demand (e.g. a certain multi-media system combined with a massage seat lead among other parts to the demand of a stronger battery). Part demand is divided in high variant parts and low variant parts.

In a matrix production system, the production is performed in workstations that are distributed on the shop floor. In contrast to function orientation of the classical job shop production, the matrix production is characterized by variant mix orientation. This means that the workstations in matrix production are arranged in a flow-oriented manner, enabling short material flows. In addition, each workstation can perform one to several productions step that enables to produce different product variants (Nyhuis et al. 2019), each in need of certain parts and resources (workers, machines, and tools). The assignment of production steps to workstations is an optimization problem itself and considered as given for the scheduling task. The material flow is divided in the main product flow (e.g. the car chassis or the bicycle frame) and the part flow from the warehouses or supermarkets to the stations. The main product is positioned on an AGV at the beginning of the process creating a fixed link between them that is only separated after all production steps have been completed. The AGV has to constantly choose the next production step based on an assembly priority graph (defined by the order) and the available capacity of workstations and parts. High variant parts cannot be stored at the workstations for all products because of their variety while low variants parts can be stored in a buffer at the station normally refilled by a Kanban system. Effectively the supply of high variants parts is the critical logistic process.

Filz et al. (2019) compared three different part supply approaches (shopping basket, tugger train and AGVs) in a simulation study. In the part supply using shopping basket (1), parts are already attached to the main part AGV and refilled or exchanged at a logistics station, if the capacity is insufficient for all parts needed in the production process. The tugger train (2) supplies the workstation buffers continuously in its routes and direct deliveries (3) is based on a direct AGV-based part transport to the station. The part supply through direct deliveries by AGVs performs best which is also the approach used by most of the papers and consequently used in this one.

The scheduling consists of assigning production steps and workstations to product AGVs based on the available capacities and parts as well as on the performable next production steps given by the assembly priority graph. The aim is to minimize the average lead time of the orders and to maximize the average capacity utilization of the workstations. Relevant configurable parameters are the number of AGV for part supply as well as the number of product AGVs that define the number of orders simultaneous being processed by the system. In contrast to classical job shop scheduling, it is possible that a job can be executed on different workstations. Additionally, multiple resources can participate in one process execution. Hence, the problem can be classified as multi resources flexible job shop problem (Zhang et al. 2019).

Since matrix production is meant to outperform flow shop systems especially in the case of short time disruptions of the planned capacities, scenarios comprise workstation breakdowns and delay of part availability.

3 LITERATURE REVIEW

Several papers have been published in the field of matrix or modular production systems in recent years that range from prototypical demonstrators, e.g. the SMART FACE demonstrator (Blesing et. al. 2017), to simulation studies (Schönemann et al. 2015). Since line production currently predominates in modern industry, it is used as a basis for comparison (Perwitz et al. 2022). The papers show a better suitability of the matrix production system for the arising requirements of multi-variant production programs. Since the results seem promising, they are continuously improved and validated. Research is divided into two main areas. In production system design, Bergmann (2022) defines the structure of the production system while production control development uses the resulting structure to deal with the challenging task of finding efficient scheduling approaches for controlling the production flow in this highly flexible environment (Hofmann et al. 2019; Perwitz et al. 2022; Mayer et al. 2021). In general, it was shown that multi agent-based systems are more suitable for the production control in these flexible production systems than central approaches (Feldkamp et al. 2019).

Synchronization of main product and part flow is one of the most critical aspects in scheduling. If parts do not reach the workstations before the product, the workstations stay idle and the main product flow stops. Unfortunately, quantities, times, and locations of multi-variant part demands cannot be known far in advance without losing flexibility in the main product flow (Skubowius et al. 2019). Nevertheless, it can be stated that most papers focus only on controlling the main product flow, leaving a gap considering in-plant part supply. The few approaches dealing with part flow at all can be clustered into two groups. Papers of the first group reducing flexibility of the main product flow to predict future part demand in advance. The challenge of unknown part demands was faced by simulative forecasting of part demands (Skubowius et al. 2019), where the workstations can be pre-stocked with parts. While this approach increases efficiency it reduces the flexibility of the control system in case of disruptions. Mueller et al. (2021) handle the problem of unknown demands with frozen periods. Which also decreases the ability to deal with disturbances caused for example by machine breakdowns in the production control but avoids the ad hoc part requests.

The second group decouples the part flow from the main product flow, assuming that in-plant logistics will manage the in-time part supply and thus schedule the main product processes independently. Filz et al. (2019) add part transportation orders to a queue, after the next production process is chosen assuming that parts and transport capacities are available to supply the workstation in time. The production logistics takes the orders from that queue and transports the parts to the next workstation. Using direct delivery with AGVs this approach reaches an overall capacity utilization rate of 54%.

Schmidtke et al. (2021) references to a heuristic based approach for the part supply, described in Bányai et al. (2019), where the supply is optimized in two stages. The part supply routes are clustered first based on a transport order queue and afterwards optimized. But also, this paper cannot present higher capacity utilization for the production. Based on these results the approach was optimized among other aspects by workstations that can execute different processes and the use of two further part supply strategies (Schmidtke et al. 2023). In the first strategy, the workstations are supplied with direct deliveries triggered when a main part AGV joins the queue in front of the workstations for assembly. In the second strategy, the main part AGV takes a shopping basket with parts for three workstations when entering the production system, extended by Kanban. Large and individual parts are delivered directly. The capacity utilization of the production system cannot be risen significantly, but station failures are considered, and can be managed quite good. Still main product flow and part flow are decoupled by a transport order queue.

Summarizing the literature review, it can be stated, that if part supply is even considered it is either gained by limiting main product flow flexibility or decoupled from the main part flow. Which means, that disturbances and restrictions in part supply will most probably lead to waiting times and thus bad capacity utilization at workstations.

4 INTEGRATED SCHEDULING OF PRODUCT AND PART FLOW

Addressing the research gap derived in section 3 we introduce our integrated scheduling approach. The integrated production control and scheduling of the main product and part flow in this paper is implemented as a multi-agent system based on the architecture presented by Blesing et al. (2017). The communication is based on the Contract Net (CNET) protocol, used by most of the papers, including Blesing et al. (2017).

The agent types and their interaction with each other as well as the sub task they are responsible for to control the production system are represented in Figure 1. The first agent type is the order management agent (A.1). It is responsible for all currently available unreleased orders, and the order release sequence. The order agent (A.2) is responsible for the completion of its current order in interaction with the resource agents (A.3) that are responsible for one or more resources and the processes executed by these resources. In general, the resource agent can administrate workstations warehouses/ supermarkets and vehicles; sub resources such as the buffers of a workstation are also included. To create a common production schedule, the coordinator agent (A.4) as the central instance is responsible for the scheduling of the decentral specified partial schedules of the resource agents during a certain process request. The last module is the simulation (A.5) that executes the planned process executions and returns the results in form of actual process executions.

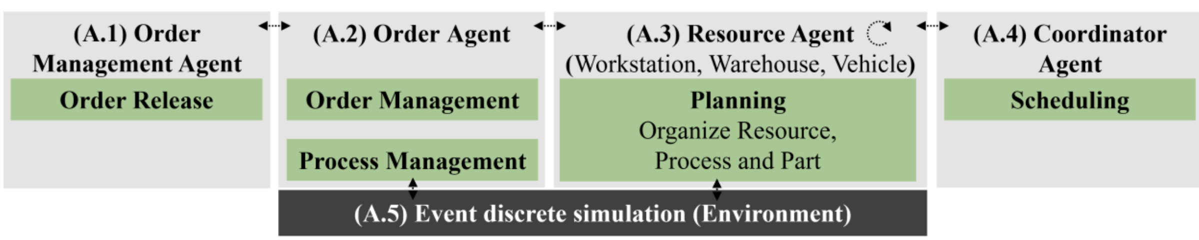

Figure 1: Agent architecture and simulation.

In the following the procedure and interaction between the agents to plan and execute process executions is presented in more detail (Figure 2). The steps are executed by the agents described above. In the following, each step is explained in more detail beginning with the first step "Order Release".

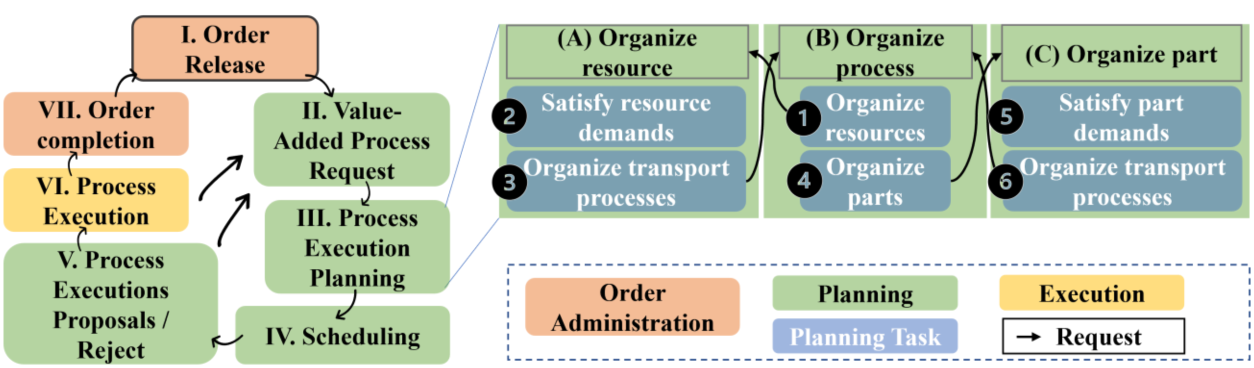

Figure 2: Production control and scheduling procedure.

4.1 Order Release

Different to Blesing et al. (2017) the order agent requests an order from the order management agent, if it does not have any order to process (Pull) and not vice versa (Push), assuring that always the same number of orders (equal to number of order agents) is in the system based on Constant Work In Process (ConWIP) (Kuhn et al. 2008). If the number of orders in system is required to be variable, this could be easily implemented as another behavior of the order management agent. Hence, the amount of main product AGVs determine the work in process (WIP). After the order management agent releases (step I) a new order, the order agent requests an AGV that is linked to the order until the order is completed.

4.2 Value Added Process Request

Afterwards a value-added process is selected, based on the possible set of processes given by the assembly priority graph. Therefore, a process is selected from the possible set of processes randomly but weighted by the parts needed, assuming that the more parts needed for the process execution the more complex the scheduling, because more sub processes must be scheduled. Accordingly, processes with more part demand are most probably scheduled first. Afterwards, the execution of the selected value-added process is requested from the possible process providers represented by resource agents. The request is valid for a defined time window and includes a preference realized as a mathematical function valid for the requested time window. If a request fails, the time window is extended, and the probability is increased that a process with less part demand is chosen.

4.3 Process Execution Planning

Process providers are always the resource agents that are responsible for the main resources which have the capabilities to perform the process (e.g. an assembly station). These resource agents are responsible for “Process Execution Planning” that is shown in more detail in Figure 2 (top right). The agents receive the process requests (A) and try to compose a feasible planned process execution. Therefore, firstly resources (1) needed are requested from the resource agents responsible for them to participate in the process execution. If a resource is requested (B), the concrete availability within the time window is evaluated (2). Available resources found are offered to the requester, if not the request is rejected. If the requested resource is not already available at the location of demand, the needed transport processes are determined and requested (3). After the resources are organized, the parts are requested (4) with “Organize parts” (C) quite similar to the “Organize resource” algorithm: the availability of a part is evaluated (5) and if the part is available the transport processes are requested if needed too (6). After each request, the results are combined.

When all requests are processed, respectively all possibilities for process executions are determined, each of them forms an AND-OR-tree (see Figure 3). The tree’s root node is the process execution requested by the order agent to execute the process on a certain workstation. Each node represents either a process, a resource or a part that is needed to make the parent node available. Each node also includes the availability over time of the corresponding item. OR-nodes need one of the child nodes while AND-nodes need all its childes to be made available. Each participating resource agent sends its AND-OR tree to the central coordinator agent to find a good and feasible solution among all requests.

4.4 Scheduling

The coordinator agents’ task (step IV “Scheduling”) is to solve the local optimization problem to find the best schedule for the current round of requests. In the current version of our implementation this is done by a search heuristic that tries to find a feasible solution based on application experts knowledge. The coordinator agent takes the AND-OR trees of all the requests and combines them in the Job Pool (Figure 3). Thus, the Job Pool contains all requests from the order agents. Afterwards a request sequence is created using a priority function (3). At the beginning, only the root nodes form the frontier, which is a set of frontier nodes s_n where $n \in \mathrm{N}$ is the id of a frontier node. Starting the procedure, the request sequence is created by opening the frontier node that has the best evaluated sub tree based on the preference function and by searching for a feasible solution for it. OR nodes and AND nodes are extended based on the same priority function (3) until a leaf node is found.

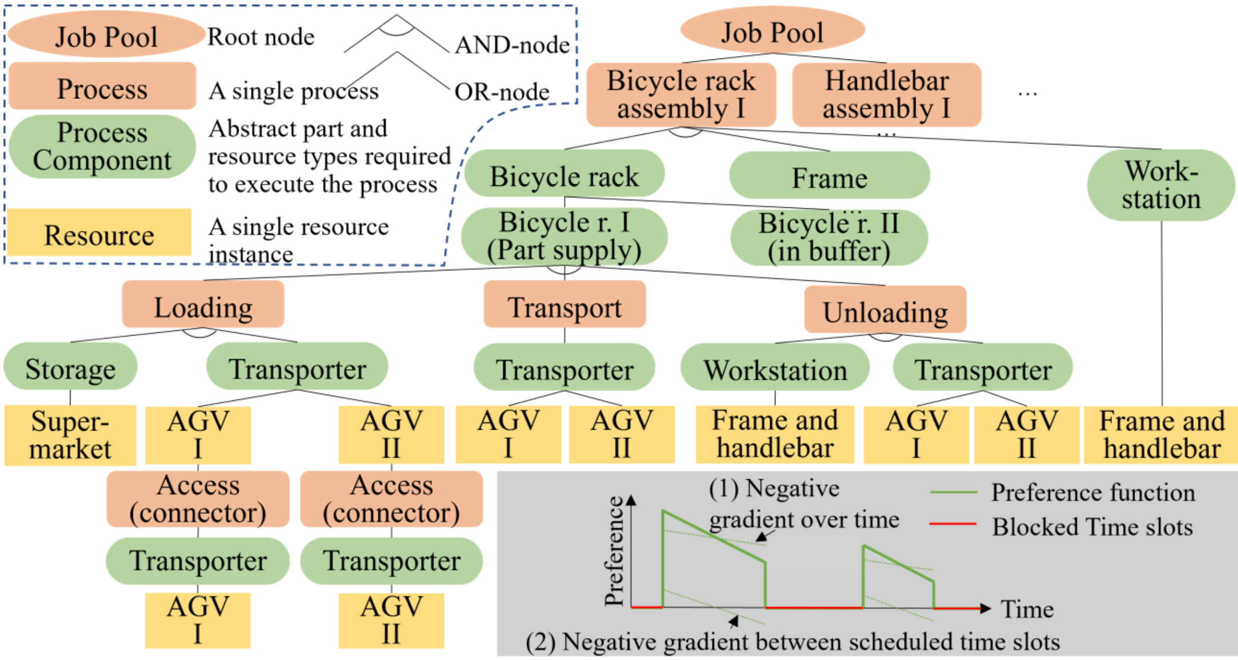

Figure 3: Scheduling AND-OR Tree example case with preference functions.

The priority function (3) is intended to ensure a fast runtime and high resource capacity utilization. In accordance the influencing factors $s_{x,n}$ of the priority function are the number of predecessor processes $s_{pp,n}$, processes that must be executed before the process as precondition. Furthermore, the earliest preferred start time $s_{e,n}$, the average capacity utilization $s_{c,n}$ of all resources in the corresponding sub tree, the processes a resource is participating in $s_{pr,n}$ and the reachable leaf nodes $s_{r,n}$ (2). The number of predecessors is used to ensure that the processes are scheduled in the correct chronological order and less replanning is needed. The earliest start time and the capacity utilization of the considered period (continuously updated) is used to ensure that waiting time in between the jobs is minimized. Furthermore, $s_{pr,n}$ is used to give the other requests the highest possible chance to get an available resource and $s_{r,n}$ to prefer smaller sub trees and therefore parts, that do not need any supply processes because they are already available at the demand location. The weights in (2) are estimated by testing different parameters settings with the aim to avoid replanning and to find a fast solution which does not claim to be optimal. In general, the values in (2) are calculated with application of (1) based on the frontier node itself and their children. Therefore, the frontier node s_n with the smallest value for an influencing factor $s_{x,n} \in \{s_{pp,n}, s_{e,n}, s_{c,n}, s_{pr,n}, s_{r,n}\}$ out of all values $S_{x,m}, m \in N$ is assigned the value 1, otherwise 0.

$$fmin\,(s_{x,n}) = \begin{cases} 1 \text{ if } s_{x,n} = min_{m \in N} S_{x,m} \\ \quad 0\ else \end{cases} \tag{1}$$

$$g(s_n) = fmin(s_{pp,n}) * 13 + fmin(s_{e,n}) * 6 + fmin(s_{c,n}) * 4 + fmin(s_{pr,n}) * 1 + fmin(s_{r,n}) * 1 \tag{2}$$

$$argmax_{n \in N}(g(s_n)) \tag{3}$$

In the scheduling itself, a preference function of the process and the preference functions of the resources involved are added to form a combined preference function. The time slot with the highest preference value (integral) based on the combined preference function is selected. The aim of the scheduling procedure is the maximization of the resource capacity utilization. To achieve this, waiting times should be avoided. Accordingly, time slots earlier in the future should be preferred. To achieve the intended behavior, preference functions (see bottom right on Figure 3) of the resources are modeled by linear functions with a negative gradient interrupted through already blocked (scheduled) periods. In addition, short unscheduled time windows should be avoided because these can remain unscheduled in further scheduling procedure due to their shortness. Therefore, the values are raised at the beginning of each unscheduled window.

On the way back to the root node the schedule is created step by step by adding all children of AND nodes and one children of the OR nodes. In the AND nodes compatibility to the already scheduled processes is ensured (e.g. that no resource is used at the same time for different process paths). If a feasible solution for one request can be found, the schedule created for a tree of the Job Pool is confirmed. Hence, parts and resources are blocked, and availability is updated for all other requests. The results are sent back to the planning resource agents.

4.5 Process Execution Proposal/ Reject, Process Execution and Order Completion

The planning resource agents take the results from the coordinator agent and pass them on to their requesters. This is repeated until the order agent either receives process execution proposals (V.) also containing process executions for the part supply etc. or a rejection message if the request failed. In the latter case, execution step VI is skipped and starts again with step II. If several process execution proposals are available for the order agent, they are evaluated, and the best proposal is accepted. The resource agents of the main resource(s) are afterwards responsible to forward the process executions associated with accepted proposal (and sub proposals) to the event discrete simulation (environment) for execution. The order agent waits until the simulation returns the results in form of actual value-added process-execution (VII.). Based on the result the order agent chooses the next value-added process (II.) and repeats the procedure. If all value-added processes required are finished, the order can be closed (VII.) and the link with the main part AGV can be reset. Afterwards a new order is requested (I.) from the order management agent and the process starts over again.

5 APPLICATION CASE: INDIVIDUALIZED BICYCLE ASSEMBLY

To show the relevance of integrated scheduling, a bicycle assembly reference model was designed and validated in a continuous dialogue with an expert from the automotive sector. It provides the possibility to compare the integrated part flow scheduling with the decoupled part flow scheduling in a simulation. The bicycle assembly contains five workstations with a processing capacity of one bicycle at a time. The shop floor layout is presented in Figure 4 (a). Almost all of them are standardized, meaning on each workstation several different assembly processes can be executed. The only exception is the painting station, which can only perform the painting process. The painting station is furthermore a bottleneck in the assembly priority graph as shown in Figure 4 (b), since no process can be executed alternatively at this point of the priority graph. It is assumed that the painting process time is approximately five times higher than the average of all assembly processes. The queue for main part AGVs waiting to enter a workstation is not limited. Each workstation is equipped with a part buffer, where the low and high variant parts can be stored, assuming no capacity limitation. The high variant parts are delivered by high variant part AGVs from the supermarket to the workstations. In addition, the main part AGVs transport the main products (frame). Each AGV has a capacity of one part and the same speed. Acceleration (e.g. start-up) is not considered. It is ensured that two AGVs cannot load or unload at the same station (supermarket or buffer) at the same time but the AGVs can pass each other and other resources to reach their destination, which means that potential traffic jams are not considered. The supermarket always stores sufficient parts to meet the demand of the processes.

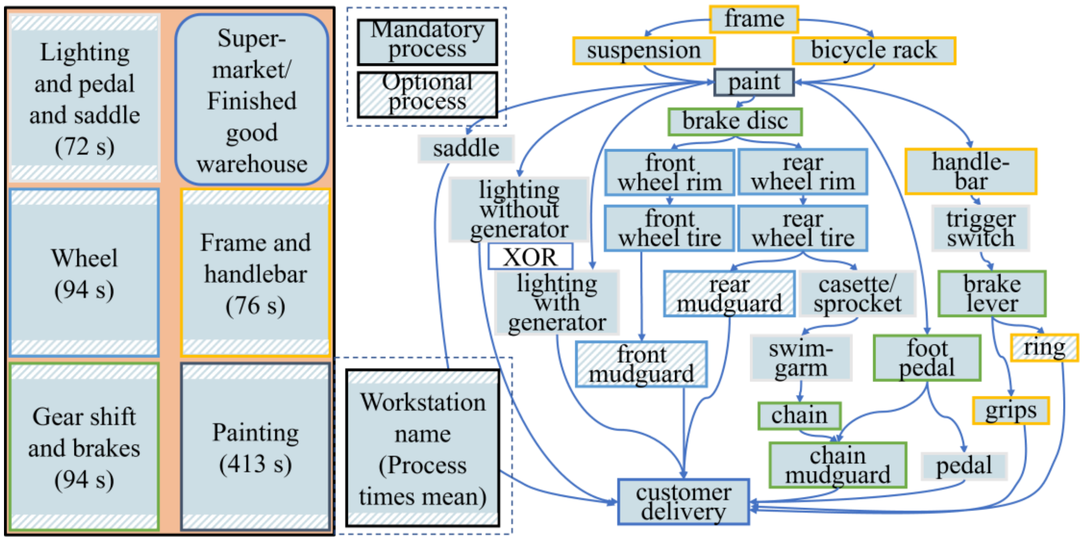

Figure 4: (a) Shop floor layout; (b) Assembly priority graph.

To represent multi-variant product orders, the products are randomly generated by choosing one feature from each of 11 feature clusters. Most features need one or more processes that must be executed, to create the product. On this basis, an order can choose one of nine bicycle frames available, to which some of the 141 possible parts are added in 17 to 22 value-added processes resulting in approximately 600 million different product configurations. As seen on the priority graph in Figure 4 (b), optional processes for example the ring or the mudguard assembly are included. Furthermore, the lighting assembly process is available in two different constellations. The one with a generator has a second predecessor and the other without did not need them. Additionally, the customer delivery process is performed from the finished goods warehouse. The process design was developed to be as close as possible to a real bicycle production. Moreover, care was taken to ensure a high degree of flexibility in using the potential of matrix production. The processes that can be executed by the same workstation are presented in the same color in Figure 4 (b). For example, the first process "frame assembly" (orange color) is performed by the same workstation as "handlebar assembly". Each process can only be performed by one workstation. As stated in (Schmidtke et al. 2023) the process times of the assembly can be different in the matrix production. The process times for the part supply differ between 20 and 49 seconds for one way between a workstation and the supermarket. Therefore, a complete part supply including accessing the supermarket and loading the part takes on average about 60 seconds in a simulation run. In comparison the mean process time for a value-added process executed on workstations is 98 seconds. In contrast to realistic behavior, all process times are assumed to be deterministic. This is a limitation that does not affect the results because replanning is not considered. To be close to reality, however, future research could also include stochastic behavior through normal distributed process times often used in this context (Schmidtke et. al. 2023; Filz et al. 2019).

6 EXPERIMENTS AND RESULTS

To validate the relevance of integrated scheduling it is compared with a decoupled scheduling of the part supply. For the decoupled scheduling approach, the product flow is scheduled first and then the part supply required for the scheduled processes is handled afterwards. For this purpose, a transportation order queue is used, which contains all required parts including their required quantities, times, and locations. Based on the queue, the part supply AGVs are scheduled according to the first in first out (FiFo) principle.

Before the comparison of both scheduling approaches, in the *first scenario – reference scenario (S1)* the WIP is varied without considering the part supply to determine the most promising values for the WIP and gain a reference value for the following evaluation of the applied scenarios. The order agents can plan with a time window of ten minutes in the future. The value was determined in consideration of the negotiation message traffic and the probability of a negotiation failing. If the negotiation fails, the value-added process selection is repeated with increased probabilities to choose a value-added process with a minor part demand by which a higher capacity utilization of the workstations is to be achieved. Moreover, the planning horizon is increased by two minutes to ensure that most processes can be scheduled within this additional time but if not other processes with less material demands get the chance to be scheduled. The WIP (thus the number of main product AGVs) is raised from six up to thirty with step size two and step size one in the most promising areas. The number of orders is set to two hundred in each simulation run and the simulation ends when all orders are processed. Considered are the capacity utilization of the workstations (later also for the logistic system) and the lead times of the orders. To ensure that the KPI values do not include the unsettled part of the simulation, only the middle of the simulation where capacity utilization is neither increasing nor decreasing constantly is considered. To identify this settled part of the simulation, the mean value of a 500-minute time window in the middle of the simulation is set as threshold. The 30-minute passing at the beginning and at the end forms the observation period. The values are set based on a visual inspection of the graphs. Within this time window only completely processed orders are considered.

Experimental results are illustrated in the graphs represented in Figure 5. The capacity utilization rises in negative exponential manner. The workstation "wheel" has reached a max value of 100 % with a WIP of 19. This is also true for the overall production system capacity utilization. Although overall capacity utilization may still rise slightly, the increase is smaller. The lead time of the orders increases linearly. Hence, we decided to evaluate the scenarios with a ConWIP of 19 orders, where the capacity utilization is 88.6 % and the lead time 158 minutes.

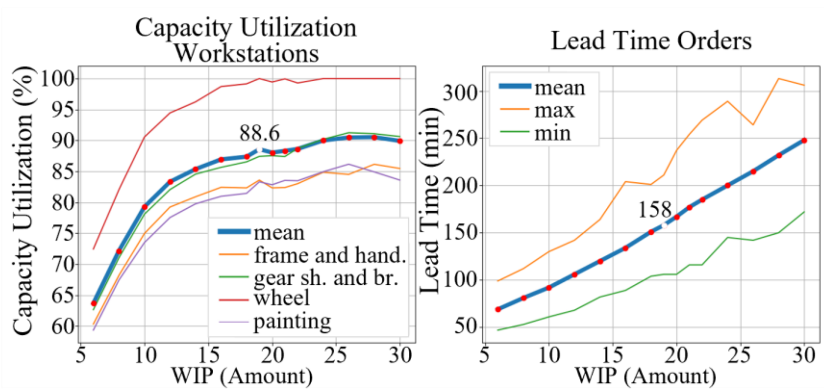

Figure 5: Reference values of capacity utilization (a) and lead time (b).

In the following scenarios, the part supply is considered. For the sake of simplicity, the supply of low-variant parts is not considered, and it is assumed that the parts are always available at the workstations using a Kanban system. This abstraction is valid since supplying these parts is no big challenge. The high variant parts are supplied through direct deliveries with AGVs. In the *second scenario – decoupled scheduling scenario (S2)* the part supply is scheduled after the main product plan is fixed. Hence, the main part AGVs must wait at the workstation until all required part supplies have arrived at the workstation. This also affects the subsequent process executions on the workstation. The *third scenario – integrated scheduling scenario (S3)* schedules the part supply within the scheduling of the product flow. If a part or a transporting part AGV is not available, the main product cannot be scheduled in the requested time window and another option has to be found.

To show the effect of integrated part flow scheduling, it is investigated how the two approaches perform when the size of part supply AGV fleet for high variant parts changes. Especially in case of bottlenecks the integrated approach is believed to perform better. A further parameter influencing the performance of the part supply, is the ratio of part supply lead times compared to the assembly lead times. Hence, as a second parameter, the speed of the high variant part AGV is varied. For the first comparison, the AGV fleet size is varied from two to six AGVs, as seen on Figure 6. For three AGVs the workstation capacity utilization for *S2* is approximately 5 % higher and the lead time 13 minutes lower. As expected, the integrated scheduling approach performs better, but when the number of AGVs rises, the KPIs of both scenarios level, because sufficient AGVs are available to meet all the part supply demands. Also, the differences between four and six AGVs are not significant. The capacity utilization with four AGVs is with 87 % (*S3*) only 1.6 % under the target capacity utilization from *S1* and can only be slightly improved with more AGVs.

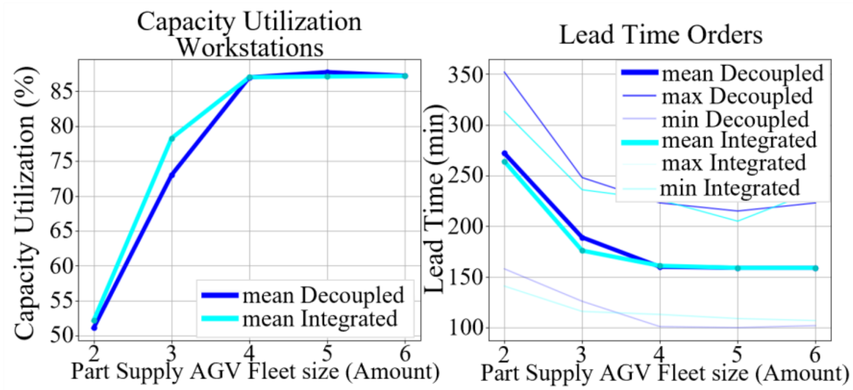

Figure 6: Comparison of S2 (blue) and S3 (red) for capacity utilization (a) and lead time (b).

The findings can be confirmed through the decrease of speed of the high variant part AGVs. Therefore, the speed is halved from 1 m/s to 0.5 m/s so that the lead time of a single part supply is almost doubled. The results are similar and shown in Figure 7. Whereby 87 % capacity utilization is reached with six AGVs seen in Figure 7 (a). The maximal capacity utilization difference can be increased slightly for the AGV fleet size of four, probably because it becomes a more relevant bottleneck with longer transportation times. Moreover, the capacity utilization of the high variant part AGV fleet is considered in Figure 7 (b). Although parts cannot be delivered in time to achieve high-capacity utilization of the workstations, the fleet seems to be underutilized in *S2 (A).* In contrast, the capacity utilization rate of *S3 (A)* is at 100 % for an AGV fleet sizes three and four. In general, however, the capacity utilization rates are at a high level for both scenarios and decreases with increasing fleet size. Additionally, there is a strong correlation between the capacity utilization of the part supply fleet and those of the workstations in case of the part supply being a bottleneck.

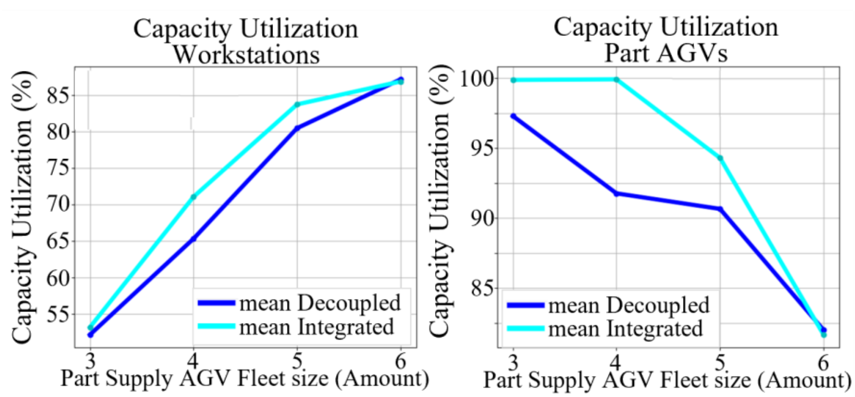

Figure 7: Comparison of *S2* (A) (blue) and *S3* (A) (red) for the capacity utilization of the workstations (a) and the high variant part AGV (b).

The experiment results show that the integrated scheduling can lead to better performance of the production system because the dependency of the main product flow from the part supply can be considered within the production scheduling. This has an impact on both the capacity utilization of workstations and the throughput time of orders. The effect is salient, as shown in Table 1, when the capacity utilization of the part AGV fleet is an optimization criterion and the part AGVs become a relevant bottleneck. In the implemented scenario, the capacity utilization could be increased by 5.2 (5.6) % and the order lead time could be decreased by 6.9 (8.5) %.

Table 1: Results of experiments in the bottleneck case.

	S2	S3	Diff.	S2 (A)	S3 (A)	Diff.
Number of part supply AGVs	3			4		
Speed of part supply AGVs	1 m/s			0.5 m/s		
Capacity Util. Workstations [%]	73.0	78.2	5.2	65.4	71.0	5.6
Capacity Util. Part AGVs [%]	92.5	99.2	6.8	91.8	99.9	8.1
Lead Time Orders [min.]	189	176	13	212	194	18

7 DISCUSSION AND OUTLOOK

In this paper we showed that integration of part flow scheduling is promising to increase capacity utilization of workstation and the transport fleet as well as to reduce order lead times, especially in the case of bottlenecks and disruption in the supply process. This is possible if processes with parts supply bottlenecks can be replaced by processes with fewer part supply demands and therefore less workload for the parts supply. The advantage could additionally increase when traffic jams on the shop floor are considered, which we belief is a very relevant topic not addressed so far in the academic research for matrix production systems.

As outlook two areas should be addressed in the future. Firstly, the heuristic-based control approach presented in this paper should be improved and replaced by more systematic optimization procedures e.g. based on metaheuristics. Additionally, further decentralization by introducing real prize-based negotiations between the agents could be interesting especially when scheduling will additionally include resources from different companies in the future. In such a decentral scenario reinforcement learning approaches could also be tested to improve the agents' behaviors. Secondly, we consider the presented bicycle application case including model and simulation as a base for further research. Until now the scientific community in the field lacks of a publicly available use-case to benchmark different approaches and make results comparable. Our aim is to make the presented environment available for other research. Nevertheless, some improvements have to be implemented beforehand. A better balancing of lead times and stations capabilities as well as including detailed traffic behavior in the simulation, to investigate spatial bottlenecks.

ACKNOWLEDGMENTS

This research was supported by the Excellence Center for Logistics and IT funded by the Fraunhofer-Gesellschaft and the Ministry of Culture and Science of the German State of North Rhine-Westphalia.

REFERENCES

Bányai, Á., I. Béla, E. Glistau, N. I. C. Machado, P. Tamás, F. Manzoor and T. Bányai. 2019. "Smart Cyber-Physical Manufacturing: Extended and Real-Time Optimization of Logistics Resources in Matrix Production". *Applied Sciences*. 9(7).

Bergmann, S. 2022. "Optimization of the Design of Modular Production Systems". In *Proceedings of the 2022 Winter Simulation Conference*, edited by B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, 1783-1793, Piscataway: Institute of Electrical and Electronics Engineers.

Blesing, C., D. Luensch, J. Stenzel, and B. Korth. 2017. "Concept of a Multi-agent Based Decentralized Production System for the Automotive Industry". In *Advances in Practical Applications of Cyber-Physical Multi-Agent Systems*, edited by Y. Demazeau P. Davidsson, J. Bajo, and Z. Vale, 19-30, Cham: Springer.

Filz, M., A. Herrmann, J. Gerberding, C. Herrmann, and S. Thiede. 2019. "Analyzing Different Material Supply Strategies in Matrix-structured Manufacturing Systems" *Proceda CIRP 81*:1004–1009.

Feldkamp, N., S. Bergmann, and S. Straßburger. 2019. "Modelling and Simulation of Modular Production Systems". In *Simulation in Produktion und Logistik 2019*, edited by M. Putz & A. Schlegel, 391-401, Auerbach: Wissenschaftliche Scripten.

Göppert, A., G. Hüttemann, S. Jung, D. Grunert, and R. Schmitt. 2018. "Frei verkettete Montagesysteme". *Zeitschrift für wirtschaftlichen Fabrikbetrieb* 113(3):151-155.

Hofmann, C., N. Brakemeier, C. Krahe, N. Stricker and G. Lanza. 2018. "The Impact of Routing and Operation Flexibility on the Performance of Matrix Production Compared to a Production Line". In *Advances in Production Research Proceedings of the 8th Congress of the German Academic Association for Production Technology (WGP),* edited by R.Schmitt, abd G. Schuh, 155–165, Cham: Springer.

Greschke, P. 2016. *Matrix Produktion. Konzept einer taktunabhängigen Fließfertigung*. Norderstedt, Books on Demand.

Kern, W.; Rusitschka, F.; Kopytynski, W.; Keckl, S.; & Bauernhansl, T. 2015. "Alternatives to Assembly Line Production in the Automotive Industry". In *Proceedings of the 23rd international conference on production research (IFPR)*. August 1th-4th, Manila, Philippines.

Kuhn, A., P. Nyhuis, and M. Schmidt. 2008. "Logistikprozesse in Industrie und Handel". In *Handbuch Logistik*, edited by D. Arnold, H. Isermann, A. Kuhn, H. Tempelmeier, K. Furmans, 295-370, Berlin, Heidelberg: VDI Buch, Springer.

Lempp, M. and P. Siegfried. 2022. "Characterization of the Automotive Industry". In *Automotive Disruption and the Urban Mobility Revolution*, edited by Martin Lempp und Patrick Siegfried, 7–24. Cham: Springer International Publishing (Business Guides on the Go).

Mayer, S., D. Gankin, C. Arnet, and C. Endisch. 2019. "Adaptive Production Control with Negotiating Agents in Modular Assembly Systems". In *2019 IEEE International Conference on Systems, Man and Cybernetics (IEE SMC)*. Bari, Italy, October 6 th-9th, 120–127.

Mayer, S., Classen, T., and C. Endisch. 2021. "Modular Production Control using Deep Reinforcement Learning: Proximal Policy Optimization". *Journal of Intelligent Manufacturing*. 32:2335‑2351.

Mueller, D., and C. Ganß. 2021. "A Simulative Study of an In-plant Frozen Period for Demand-Based Material Supply in Matrix Manufacturing Systems". In *2021 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. Singapore, Singapore, December 13th-16th, 185–189.

Perwitz, J., T.; Sobottka, N.-J. Beicher, A. Gaal. 2022. "Simulation-based evaluation of Performance Benefits from Flexibility in Assembly Systems and Matrix Production", Procedia CIRP, 107:693-698.

Skubowius, E., H. Bayhan, C. Schwede, and M. ten Hompel. 2019. "Simulation als Voraussetzung zur dezentral gesteuerten Materialbereitstellung in cyberphysischen Produktionssystemen", In *28. Deutscher Materialfluss-Kongress 2019*, edited by VDI Wissensforum GmbH, 77-90, Düsseldorf: VDI Verlag.

Schmidtke, N., A. Rettmann, and F. Behrendt. 2021. "Matrix Production Systems - Requirements and Influences on Logistics Planning for Decentralized Production Structures". In *Proceedings of the 54th Annual Hawaii International Conference on System Sciences*, edited by T. X. Bui, Honolulu: Univ. of Hawaii at Manoa, 1665-1674.

Schmidtke, N., A. Rettmann, J. Mohr, and F. Behrendt. 2023. "Investigation of Material Supply Strategies to Increase Resilience in Matrix Production Systems". In *Proceedings of the 56th Annual Hawaii International Conference on System Sciences*, edited by T. X. Bui, Honolulu: Univ. of Hawaii at Manoa, 1458-1467.

Schönemann, M., C. Herrmann, P. Greschke, and S. Thiede. 2015. "Simulation of Matrix-Structured Manufacturing Systems". *Journal of Manufacturing Systems* 37:104–112.

Zhang, J., G. Ding, Y. Zou, Q. Shengfeng, J. Fu. 2019. "Review of Job Shop Scheduling Research and its New Perspectives under Industry 4.0". *Journal of Intelligent Manufacturing*. 30:1809‑1830.

Nyhuis, P., N. E. Rochow, M. Krause, D. Pischke, M. Seitz, K. K. Vivian. 2021. "Organisationsformen der Produktion". In *Journal of Production Systems and Logistics* 1:1-35.

AUTHOR BIOGRAPHIES

ADRIAN FREITER is a reasearch assistant at the Fraunhofer ISST in Dortmund. Beyond that he is a research master student of University of Applied Sciences and Arts Bielefeld and studies data science with research focus on the application of multi-agent-based simulation in the production and logistics. His email address is adrian.freiter@isst.fraunhofer.de.

CHRISTIAN SCHWEDE is Professor for Big Data Analysis at the University of Applied Science and Arts Bielefeld. He is member of the Center for Applied Data Science and Board member of the Institute for Data Science Solutions. Additionally he is a senior scientist at Fraunhofer ISST in Dortmund resposible for AI in Logistics. His reseach area is the application of data- and model-driven methods to optimize logistics and production. He is specialized in automotive logistics and agent-based dezentralized production systems. His email address is christian.schwede@hsbi.de.