An Artificial Hormone-based Algorithm for Production Scheduling from the Bottom-up

Wilfried Elmenreich¹, Alexander Schnabl¹ and Melanie Schranz²

¹Institute of Networked and Embedded Systems, Alpen-Adria-Universität Klagenfurt, Klagenfurt, Austria
²Lakeside Labs, Klagenfurt, Austria

Keywords: Production Scheduling, Job-shop Scheduling, Artificial Hormone Systems, Self-organizing Systems, Swarm Systems.

Abstract: This paper presents a model for supporting a production scheduling system with an artificial hormone algorithm. The system consists of lots that have to undergo a number of processing steps on different machines. The processing steps for a lot are formalized in a recipe assigned to the lot type. Since the steps in the recipe have to be processed in order, the given system allows choice only in the context of selecting a particular machine for the next step and in changing the processing order of waiting lots at a machine. Optimization of such a job-shop scheduling system is an NP-hard problem. In the approach proposed by this paper, artificial hormone systems are used to express the urgency of a lot and the need for new lots at a machine type, thus providing a system using local information for optimization. Results indicate that the artificial hormone system provides an improvement of around 5% over a First Come-First Serve approach.

1 INTRODUCTION

The production of logic and power integrated circuits (ICs) in the semiconductor industry is a highly dynamic process (Geng, 2018). Unlike the high-volume production of memory ICs, in the logic and power sector, the wafer production has a huge product mix, dynamic changes in the system and a high number of processing steps and involved machines (Khatmi et al., 2019). Weekly workloads can involve around $10^5$ operations on $10^3$ machines (Teppan, 2018). Optimizing such a system for WIP (work in progress) and flow factor is an NP-hard problem (Garey et al., 1976), where existing dispatching rules and linear optimization methods cannot cope with the NP-hard search space (Lawler et al., 1993), and thus cannot consider the entire system behavior due to computational complexity. Therefore, they cannot exploit the optimization potential in job-shop scheduling in semiconductor production systems (Khatmi et al., 2019).

To overcome this issue, we model the production plant as a self-organizing system of agents that interact with each other in a non-linear way. Therefore, we aim at achieving near-optimal solutions in feasible computation time. Furthermore, such a system is also able to adapt to changes in the environment, scale with the number of agents, and is robust against a single point of failure, as the actions depend on local interactions (Heylighen, 2001). Moreover, local rules and local interactions allow overcoming the huge computing time of centrally performed linear optimization. One possibility for a self-organizing approach are algorithms modeled after biological hormone systems. Such artificial hormone systems are inspired by the biological endocrine system that is adjusting the metabolism of tissue cells in our body (Turing, 1952; Sobe et al., 2015). It is a class of self-organizing system with properties like scalability, adaptability, and robustness (Prehofer and Bettstetter, 2005) that can be used in networked technical applications to coordinate a set of complex agents interacting with each other (Elmenreich et al., 2009; Böszörményi et al., 2011; Sobe, 2012). Such algorithms are especially of interest for such large cyber-physical systems, as, for example, in the semiconductor industry, where traditional control or scheduling mechanisms are at their limit.

In this paper, we describe an artificial hormone algorithm that can be applied to optimize the processing of lots and to balance the load between machines of the same type. Section 2 describes the target system problem statement that forms the basis to integrate an artificial hormone-based algorithm in such a system in Section 3. Section 4 evaluates the approach...
3 ARTIFICIAL HORMONE ALGORITHM

The artificial hormone algorithm is engineered in a bottom-up approach to express the urgency of a lot and the need for incoming lots by machines. Therefore, artificial hormones are produced at machines which can diffuse through the production system along with the processing steps of lots. The lots act as swarm members that can be attracted by the hormone level of a machine. The algorithm is based upon five parameters, which are described with their respective mechanisms in the following. Table 2 depicts an ad-hoc definition of parameter values, which will be used in first simulations. The algorithm describes the calculations and decision mechanisms that affect the processing of lots. There are several degrees of freedom for a particular implementation in the field. For example, the hormone-related computing could take either at the machines or within a layer of a networked monitoring and control system.

### 3.1 Hormone Model

For each processing step, a corresponding hormone type exists. Hormones can be at any machine, also different hormone types can be at the same machine. Every simulation tick, hormones degrade exponentially at a given rate $\alpha$.

$$\text{hormone}_\text{amount} = \text{hormone}_\text{amount} \cdot (1 - \alpha) \quad (1)$$

An evaporation rate of 0 means that hormones do not degrade. The maximum value is 1, where hormones degrade immediately.

### 3.2 Machines Produce Hormone to Attract Lots

Each machine generates a hormone of its process type. Machines that perform the same process type issue the same hormone. If a machine would be able to handle multiple process types, it would issue, in equal parts, the respective multiple hormones instead. The goal of machines is to maximize their working time. Therefore, a machine would aim for pulling a sufficient number of lots into its queue. The amount of hormone produced by one machine is therefore calculated by

$$\text{hormone}_\text{output} = \frac{1}{\text{lots}_\text{in}_\text{queue} + \beta} \quad (2)$$

where $\beta$ is a smoothing factor $> 0$. 

---

Table 1: Parameters used in the simulation model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process types</td>
<td>100</td>
</tr>
<tr>
<td>Machines per process type</td>
<td>$U(2, 10)$</td>
</tr>
<tr>
<td>Machines with batch processing</td>
<td>50</td>
</tr>
<tr>
<td>Batch size</td>
<td>$U(2, 8)$</td>
</tr>
<tr>
<td>Processing time per production step</td>
<td>$X \sim \mathcal{N}(1.16, 0.32)$</td>
</tr>
<tr>
<td>Number of recipes (lot types)</td>
<td>100</td>
</tr>
<tr>
<td>Recipe length</td>
<td>$U(90, 110)$</td>
</tr>
<tr>
<td>Lots per type</td>
<td>$U(2, 10)$</td>
</tr>
<tr>
<td>PCT</td>
<td>$U(2, 10) \times \text{RPT}$</td>
</tr>
<tr>
<td>Total number of operations</td>
<td>$\approx 60000$</td>
</tr>
</tbody>
</table>
Table 2: Suggested algorithm parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>.3</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>.5</td>
</tr>
<tr>
<td>$\delta$</td>
<td>.2</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>.8</td>
</tr>
</tbody>
</table>

3.3 Machines Are Linked

A machine $A$ is upstream-linked to a machine $B$ if there exists a recipe that has a process of machine $B$ and a process of machine $A$ in subsequent steps. If each machine has exactly one process, the linking strength from this recipe is 1, otherwise, the linking strength is $2 / \text{number of supported processes}$. If the same sequence appears in other recipes, the linking strength adds up.

3.4 Hormone Diffuses Upstream

If upstream links exist, a part $\gamma$ of the hormone at a given machine diffuses upstream:

$$\text{upstream}_\text{hormone} = \text{hormone}_\text{amount} \cdot \gamma$$ (3)

$$\text{hormone}_\text{amount} = \text{hormone}_\text{amount} - \text{upstream}_\text{hormone}$$ (4)

For each machine connected upstream a proportional part of the upstream hormone is added:

$$\text{added}_\text{hormone} = \frac{\text{upstream}_\text{hormone} \cdot \text{link_strength}}{\sum \text{link_strengths}},$$ (5)

where $\text{link_strength}$ refers to the upstream link strength for the given hormone between the respective machines and $\sum \text{link_strengths}$ refers to the sum of all upstream link strengths for the given hormone that emerge from the sending machine.

3.5 Incoming Lots Diffuse Hormone

In addition to the mechanism above, incoming lots cause a part $\delta$ of the respective hormone at a given machine to diffuse upstream.

$$\text{upstream}_\text{hormone} = \text{hormone}_\text{amount} \cdot \delta$$ (6)

$$\text{hormone}_\text{amount} = \text{hormone}_\text{amount} - \text{upstream}_\text{hormone}$$ (7)

The machine where the lot came from receives the upstream hormone, this way a flow of lots can self-stabilize.

$$\text{added}_\text{hormone} = \text{upstream}_\text{hormone}$$ (8)

3.6 Lots Are Prioritized by Their Timing

A lot has a base priority calculated from its remaining RPT and its remaining PCT:

$$\text{base}_\text{priority} = \frac{\text{remaining}_\text{RPT}}{\text{remaining}_\text{PCT}}$$ (9)

3.7 Lots Are Attracted by Hormone

The priority of a lot is modified by hormones that are present at the machine where the lot is currently waiting

$$\text{attraction} = \sum h_i \cdot \epsilon^i,$$ (10)

where $h_i$ is the hormone of the process that is $i$ steps ahead in the lot’s recipe. Therefore, $h_0$ is the hormone of the current process. $\epsilon$ is a factor indicating the strength of a hormone’s influence.

The priority of a lot is then calculated with:

$$\text{priority} = \text{base}_\text{priority} \cdot \text{attraction}$$ (11)

At a machine, lots are processed based on their priority. If a machine has batch processing, all lots that fit are accepted for one batch.

4 EVALUATION

4.1 Evaluation Environment

The simulation was constructed in a NetLogo environment. NetLogo is an agent-based modeling software which includes the NetLogo programming language and an integrated development environment (Wilensky and Rand, 2015). An agent-based model is used to simulate the interactions of autonomous agents while also keeping track of the whole simulated system. Every autonomous agent makes decisions on its own based on predefined rules which later impact the outcome of the simulation. In the case of the problem statement, the agents are machines and lots. Before the simulation can start, the parameters have to be set. This can be achieved via the GUI in NetLogo or, for larger quantities, with configuration files. Then, the Setup function is called to
initialize the model. It uses the specified parameters or parses the given configuration files to generate the agents. After the machines and lots have been placed in the production plant, the lots are introduced into production as shown in Figure 1. In it the machines are shown as squares, batch machines are being visualized with additional lines. The lots are numbered based on their product type and the most recent path they took is marked with a line.

4.2 NetLogo Model

After the initialization of all entities is finished, the production starts. A production step runs as follows: Every lot moves to the queue of the production type specified in its recipe. It only does so if it is currently not being processed by a machine. The algorithm then updates the hormones and sorts the queues based on the priority of the lots as described in Section 3. Listing 1 shows a simplified version of the step function and Listing 2 and 3 outline two exemplary hormone functions.

Figure 1: Screenshot from the Netlogo simulation: After initialization the lots start being processed and move to the queues specified in their recipes.

Listing 1: NetLogo process step.

to step
  ask lots [ if not processing? [ move-to-queue ] ]
  ask machines [ produce-hormones self decay-hormones self ]
  sort-queues
  ask machines [ pick-lots-from-queue do-processing ]
end

Listing 2: NetLogo hormone production function.

to produce-hormones [ m ]
  let q (item ([m.machine_type] of machine m - 1) queue_list)
  let q_len_length [q.lotlist] of q
  ask machine m [ set m.hormone_amount m.hormone_amount + (1 / (q_len + BETA)) ]
end

Listing 3: NetLogo hormone decay function.

to decay-hormones [ m ]
  ask machine m [ set m.hormone_amount m.hormone_amount * (1 - ALPHA) ]
end

After the calculations are done, the machines pick the lots with the highest priority from their queue to be processed. When a lot finishes its recipe it leaves the production plant and writes its ProductionTime, RPT, and PCT to a file. This sequence is repeated until there are no more lots left in the production plant.

4.3 Headless Implementation

Using the technique specified above it is possible to gather large amounts of data by running multiple instances of the simulation simultaneously without a GUI in headless mode. The headless controller software was written in Python using the NL4Py package

1 This package was chosen instead of other NetLogo controller software or the NetLogo BehaviourSpace because the algorithm has a very high time complexity and as soon as it gets to simulating larger quantities of machines, lots, etc. there is a significant improvement in performance. Through the threading module multiple simulations can be executed at the same time. The Python code was augmented with an automated configuration file generator. It creates new files for each simulation run using randomly chosen values within the limits specified in Table 3. Listing 4 shows an example workspace for multi-threading.

Listing 4: Python headless workspace.

def start_workspace(config_file, parameters):
  workspace = nl4py.newNetLogoHeadlessWorkspace()
  workspace.openModel(model)
  workspace.command("Setup")
  ...
workspace.command("headless-go")

while (processing == 0):
    time.sleep(20)
    processing = workspace.report("IS_FINISHED")
ticks = str(workspace.report("ticks"))
print("Ticks: " + ticks)

4.4 Parameter Optimization

Another implementation of the algorithm was developed in Java so as to achieve a better way of finding optimal parameter values for the algorithm. This had various reasons such as performance and adaptability for future work. With the possibilities of object-oriented programming, the created simulation environment matching the problem statement is only based on 3 functions and it is now possible to simply exchange the algorithm of a production plant to compare different approaches with one another. The creation of new algorithms is also fairly easy and for data comparison, a simple first come, first served algorithm was also used.

5 SIMULATION RESULTS

The simulation results of the NetLogo model are evaluated based on average flow factor and tardiness of the lots. Table 3 details the used evaluation parameters.

\[ \text{flow factor} = \frac{\text{production time}}{\text{RPT}} \]
\[ \text{tardiness} = \text{production time} - \text{RPT} \]

Table 3: Evaluation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial RNG seed</td>
<td>58008</td>
</tr>
<tr>
<td>Number of simulation runs</td>
<td>100</td>
</tr>
</tbody>
</table>

The hormone algorithm was tested using the parameters outlined in Table 1, Table 2 and Table 3. Results are depicted in Table 4.

Table 4: Hormone algorithm average simulation results.

<table>
<thead>
<tr>
<th>Average</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to produce all lots</td>
<td>173 h</td>
</tr>
<tr>
<td>RPT</td>
<td>59 h</td>
</tr>
<tr>
<td>PCT</td>
<td>348 h</td>
</tr>
<tr>
<td>Tardiness</td>
<td>114 h</td>
</tr>
<tr>
<td>Flow Factor</td>
<td>2.93</td>
</tr>
</tbody>
</table>

A FCFS (first come, first served) algorithm was used to contrast the results. It ran with exactly the same test set and resulted in the values depicted in Table 5.

Table 5: FCFS algorithm average simulation results.

<table>
<thead>
<tr>
<th>Average</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to produce all lots</td>
<td>184 h</td>
</tr>
<tr>
<td>RPT</td>
<td>60 h</td>
</tr>
<tr>
<td>PCT</td>
<td>331 h</td>
</tr>
<tr>
<td>Tardiness</td>
<td>124 h</td>
</tr>
<tr>
<td>Flow Factor</td>
<td>3.05</td>
</tr>
</tbody>
</table>

With the used parameters the hormone-based production approach seems to perform slightly better than an FCFS algorithm; however, considering the algorithm parameters in Table 2 are ad-hoc values, there is a chance of improving the efficiency of the hormone algorithm by optimizing the parameters, which was done with the implementation in Java. While the NetLogo simulation ran on multiple randomly generated sample sets to gather a better general understanding of the speed of the algorithm, the Java version used the same configuration of machines, lots and products and just the parameters were altered. Table 6 and Table 7 show the fastest and the slowest simulation results and parameter values of the Java hormone algorithm. The algorithm with the optimized parameter set performs 9% better than the reference algorithm, however, it has to be noted that the optimization was done for a single test case. More importantly, while the two tables depict a potential for optimization of parameters, they also show that even with a bad parameter set the algorithm works slightly better than the FCFS algorithm. For reference, Table 8 shows the outcomes of the FCFS algorithm.

Table 6: Hormone algorithm parameters yielding the fastest result for the tested sample set.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.2</td>
</tr>
<tr>
<td>β</td>
<td>3.0</td>
</tr>
<tr>
<td>γ</td>
<td>0.6</td>
</tr>
<tr>
<td>δ</td>
<td>0.3</td>
</tr>
<tr>
<td>ε</td>
<td>0.7</td>
</tr>
<tr>
<td>Time to produce all lots</td>
<td>152 h</td>
</tr>
</tbody>
</table>

Table 7: Hormone algorithm parameters yielding the slowest result for the tested sample set.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.3</td>
</tr>
<tr>
<td>β</td>
<td>3.0</td>
</tr>
<tr>
<td>γ</td>
<td>0.5</td>
</tr>
<tr>
<td>δ</td>
<td>0.1</td>
</tr>
<tr>
<td>ε</td>
<td>1.0</td>
</tr>
<tr>
<td>Time to produce all lots</td>
<td>164 h</td>
</tr>
</tbody>
</table>
Table 8: FCFS Algorithm for the tested sample set.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to produce all lots</td>
<td>165 h</td>
</tr>
</tbody>
</table>

6 RELATED WORK

In a dynamic job shop, currently applied dispatching rules are based on heuristics with the disadvantage of not optimizing lot sequences. Although linear optimization methods are used to strengthen the scheduling procedure, they cannot cope with the highly complex, large, and dynamic search space (Lawler et al., 1993), mostly due to the excessive and unfeasible computation time in dynamic production plants. This leads to bottlenecks and work in progress (WIP) waves are generated. So far, no optimal solution for job shop scheduling has been developed using linear optimization that can be computed in polynomial time (Zhang et al., 2009).

In the course of this paper, we model the production plant as a self-organizing system of agents. Due to its non-linearity, it is able to produce near-optimal solutions for NP-hard problems in feasible computation times (Zhang et al., 2009; Dhiman and Kumar, 2017) by transforming the problem from finding an overall solution to defining a distributed algorithm that finds the solution from the bottom up. The related work on the application of self-organization in production scheduling builds upon the particle swarm optimization (Ghumare et al., 2015), and artificial bee colony algorithm (Zhang et al., 2013), and ant algorithm (Udomsakdigool and Kachitvichyanukul, 2008), to name but a few. Compared to these algorithmic approaches, we do not create a swarm that operates in a solution space to a given job-shop scheduling set. Instead, we derive a bottom-up approach, where embodied agents represent physical entities in the fab, and work with local rules from which a global behavior emerges. This presents a novel approach in the application of self-organizing principles in job-shop scheduling.

This approach can be implemented with an artificial hormone algorithm. It is a bio-inspired self-organizing algorithm to produce near optimal solutions in NP-hard, highly complex and dynamic systems based on relatively simple and local rules. The algorithm was introduced by Sobe et al. (Sobe et al., 2010) and describes the inspiration by the endocrine system of higher mammals on the technical application of sharing multimedia units. The main principle describes virtual hormones that show their interest in specific interest units with their concentration density.

These hormones are created, consumed or evaporated, and forwarded by the agents in this artificial hormone system (Szkaliczki et al., 2013). Artificial hormone systems have been applied in technical systems such as task allocation (Renteln et al., 2008), synthesis of robot controller software (Hamann et al., 2010), or content delivery in dynamic networks (Schelfhout and Holvoet, 2003). Trumler et al. (Trumler et al., 2006) propose an artificial hormone system as middleware for ubiquitous computing, as for example in smart office applications. Brinkschulte et al. present an artificial hormone system for task allocation for heterogeneous processing units in a processor grid to realize self-X features (self-configuration, self-optimization, self-healing) (Brinkschulte et al., 2007; Brinkschulte et al., 2009). Sobe et al. (Sobe et al., 2015) build a middleware inspired by an artificial hormone system for search and delivery of information units. They showcase the applications of multimedia distribution at social events, and the information dissemination in smart electrical microgrids using tens of thousands of agents. The evaluation was also performed on a theoretical foundation to show that the hormone levels converge to a limit at each agent using a set of theorems on convergence conditions (Szkaliczki et al., 2016). Finally, Dong et al. (Dong et al., 2010) successfully proposed the artificial hormone algorithm for clustering in wireless sensor networks to prolong the lifetime of the sensor nodes. Further details on the origins of the artificial hormone algorithm, including the medical background of the endocrine system and its digitization, can be found in Xu and Wang (Xu and Wang, 2011).

7 CONCLUSION

The novel contribution of this paper is the application of an artificial hormone algorithm in the context of a job-shop scheduling system such as a large semiconductor fab from the bottom-up. The algorithm builds upon five principles, which are (i) machines produce hormone to attract lots, (ii) hormone diffuses process-upstream, (iii) incoming lots diffuse hormone, (iv) lots are prioritized by their timing, and (v) lots are attracted by hormone. Via these mechanisms, machines can balance their workload by pulling required lots towards them. The algorithm has been implemented and evaluated in a NetLogo simulation model. Simulation results indicate that the artificial hormone system provides an improvement of around 5% for overall production time and flow factor.

301
ACKNOWLEDGEMENTS

We would like to acknowledge Martina Umlauft for her comments on the paper and the NetLogo implementation. This work was performed in the course of project SWILT (Swarm Intelligence Layer to Control Autonomous Agents) supported by FFG – IKT der Zukunft under contract number 867530.

REFERENCES


Turing, A. M. (1952). The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society of


