Aggregation techniques for a scheduling model in the photolithography area in the semiconductor manufacturing

Jérémy Berthier[1,2], Stéphane Dauzère-Pérès[1], Claude Yugma[1], Alexandre Lima[2]

[1] École des Mines Saint-Étienne, UMR 6158 LIMOS, CMP, Department of Manufacturing Sciences and Logistics, Gardanne, France

j.berthier@emse.fr, dauzere-peres@emse.fr, yugma@emse.fr

[2] STMicroelectronics Crolles, Department of Decisional Solutions, Team of Full Automation and Simulation, Crolles, France

jeremy.berthier@st.com, alexandre.lima@st.com

Mots-clés : *Operations research, Scheduling, Multi-objective optimization, Semiconductor* .

1 Introduction

Semiconductor manufacturing includes the most complex manufacturing processes. Scheduling problems to be addressed at the operational level involve a rich set of constraints and criteria. As a result, optimization algorithms are increasingly preferred over dispatching rules, especially in complex production areas such as the photolithography area which is considered in this paper.

2 Problem description

The scheduling problem in the photolithography area consists in scheduling a set of jobs on a set of parallel photolithography machines. Each job requires an additional resource, called reticle, that can be transported from one machine to another.

Jobs Each job needs to be processed with a given priority. Few jobs have precedence constraints (without formally considering the problem as a flexible job-shop scheduling problem), and must not exceed a maximum time lag. Engaged maximum time lag constraints when starting the schedule are modelled by assigning a deadline to each job.

Photolithography machines Each photolithography machine is qualified to process a limited subset of the jobs. It can process one job at time, and may not be available before a certain date. Processing on the machine is non-preemptive and its duration depends on both the job and the machine. Finally, a machine-dependent and sequence-dependent setup time is required to start the processing of a job.

Reticles Each job requires one reticle, which is available in a single copy. Jobs are assumed to have a competitive access to reticles (otherwise, considering them would be trivial). Besides, transporting a reticle between two machines takes a sequence-dependent duration.

3 Multi-objective scheduling model

The above scheduling problem is addressed through an Integer Linear Programming (ILP) model based on the works of [1] and [2]. Time is divided into periods of equal duration. A feasible solution is represented by binary decision variables associated with each triplet (job, machine, period). Several constraints are defined to comply with the description made in the previous section. Moreover, all jobs must be scheduled before the end of the time horizon.

Several objective functions are defined based on the factory requirements. All the criteria must be considered in the optimization problem, which makes it multi-objective. Overall, three categories of criteria are studied :

— Criteria implementing relaxed operational constraints, such as the minimization of a total risk function associated with the maximum time lag constraints and deadlines ;
— Criteria to meet production targets of the manufacturing area ;
— Criteria aiming at improving the industrial efficiency.

4 Aggregation methods

The ILP problem turns out presented in Section 3 turns out to be intractable for large instances. In order to improve the solving efficiency, problem size is reduced using two aggregation methods that are proposed for the scheduling model.

4.1 Batching of similar jobs

One investigated technique consists in batching in series jobs using the same reticle. For the sake of simplicity, jobs in the same batch are selected such that :

- None of them are involved in a precedence relationship ;
- None of them are in-process ;
- No setup times are required between any sequence of jobs belonging to the batch ;
- All jobs have at least one common qualification, which means that $\cap_{j \in b} \mathcal{M}_j \neq \emptyset$.

Inside the batch, jobs are ordered using Smith's rule applied to the average total processing time of the batch over qualified machines, i.e. by ascending order of criteria $\frac{\sum_{\cap_{j \in b} \mathcal{M}_j} p_{j,m}}{\omega_j \cdot |\cap_{j \in b} \mathcal{M}_j|}$. From the scheduler perspective, each batch b is eventually treated as a lot having the following whose priority is $\omega_b \leftarrow \max_{j \in b} \omega_b$.

4.2 Increase of the time step

A second tested techniques simply consists in increasing the time step Δ, which is equivalent to changing any time-related parameter α_t as follows : $\alpha_t \leftarrow \frac{\alpha_t}{\Delta}$. Since decision variables are time-indexed, this results in dividing their number by Δ. The obtained scheduled is later corrected using the initial value of α_t.

5 Conclusions

The mathematical model and the aggregation methods will be presented in details in the conference, together with the numerical experiments showing the limits of the model and the performances of the aggregation techniques.

Références

[1] Abdoul Bitar, Stéphane Dauzère-Pérès, and Claude Yugma. Unrelated parallel machine scheduling with new criteria : Complexity and models. Computers Operations Research, page 105291, 2021.

[2] Abdoul Bitar, Stéphane Dauzère-Pérès, Claude Yugma, and Renaud Roussel. A Memetic Algorithm to Solve an Unrelated Parallel Machine Scheduling Problem with Auxiliary Resources in Semiconductor Manufacturing. Journal of Scheduling, 19(4) :367–376, 2016.