

Solving a Flowshop Scheduling Problem with Answer Set Programming: Exploiting the Problem to Reduce the Number of Combinations

Carmen Leticia García-Mata and Pedro Rafael Márquez-Gutiérrez

Tecnológico Nacional de México - Tecnológico de Chihuahua
Chihuahua, Chih., México

clgarcia; pmarquez@itch.edu.mx

Planning and scheduling have been a central theme of research in computer science. In particular, the simplicity of the theoretical approach of a no-wait flowshop scheduling problem does not allow to perceive the problem's complexity at first sight. In this paper the applicability of the Answer Set Programming language is explored for the solution of the Automated Wet-etching scheduling problem in Semiconductor Manufacturing Systems. A method based in ranges is proposed in order to reduce the huge number of combinations.

1 Introduction

A distinctive characteristic of combinatorial problems is their massive search space. This huge domain is due to the number of possible solutions that although finite, grows exponentially with the amount of data. Some typical combinatorial problems are the search for the cheapest or shortest paths, internet data packets routing, protein structure prediction, and planning and scheduling of resources.

In theory it is possible to find the optimal solution for each combinatorial problem by conducting an exhaustive search. However, in practice finding an optimal solution is often an intractable problem, even for problems of modest size.

In this paper, Answer Set Programming (ASP) is used to explore how to solve the scheduling problem for an Automated Wet-etch Station (AWS) of a Semiconductor Manufacturing System where the optimization objective is the makespan. If a robot is not used to transfer jobs between baths, the problem can be approximated as a special case of the most general no-wait scheduling flowshop problem. A flowshop is a multi-stage production process where all jobs must pass through the same stages. There is a set J of jobs with $|J| \in \mathbb{N}$ jobs in total. All jobs must be processed sequentially, starting on machine one and following a chain of m operations O_j, \dots, O_{jm} . The operation O_{ji} must be processed for t_{ji} time units at each stage. Each job is processed at most by one machine and each machine executes at most one job. Each operation is performed without interruption in the machine that was assigned. The no-wait constraint in scheduling problems occurs when two consecutive operations of a job must be processed without interruptions. AWS's wafer etching consists in processing n jobs through a series of chemical and water baths alternately ordered. Processing in chemical baths should immediately follow the previous operation because an over exposition in the chemical bath can damage the circuit. Once the wafer is processed in a water bath, it can be used as local storage if needed.

According to Graham's notation, this problem belongs to the $F_m/no - wait/C_{max}$ class. A good survey about no-wait flowshop scheduling problems is the one by Allahverdi [1].

The three most popular AI methods used on the flowshop and AWS scheduling problem are: a) meta-heuristics, like genetic algorithms (GA); particle swarm optimization (PSO); ant colony optimization

(ACO); Tabu Search (TS); Simulated Annealing (SA); b) statistics: Monte Carlo methods and neural networks; and c) rule-based methods: CP.

One early paper reporting a solution to the AWS problem is Geiger's et al. [3]. They used the meta-heuristic TS with two additional heuristics to minimize the C_{max} : (a) Profile Fitting (PF) heuristic and (b) Nawaz-Enscore-Ham (NEH) heuristic. That paper reported better results when it is used NEH-TS. On the other hand, Bhushan and Karimi [7] proposed a solution based on TS and SA, combined with new specific algorithms for the same problem.

By contrast, there are only a handful of researchers reporting optimized solutions to the AWS scheduling problem via a combinatorial approach. One of them is due to Zeballos et al. [8] who reported a solution using CP combined with different search strategies. Similarly, Novas and Henning [5] reported another solution to the problem using CP, generalizing the problem by an innovative rolling horizon methodology and taking into account the robot's empty movements.

ASP is a language for knowledge representation, based on non-monotonic logic and designed to solve combinatorial problems [11]. ASP has been used to solve combinatorial knowledge-based problems in many different areas. Among the most outstanding applications is the decision-making system of a spacecraft [9], planning for the generation of teams in a seaport [4], and autonomous vehicles in car assembly [10].

Despite the amazing characteristics of ASP, its grounding step represents an important bottleneck in the solution process and we had to find a way to reduce the number of combinations. To this end, we use the approach of exploiting information relative to the problem structure. In particular, we propose a novel method based on ranges to reduce the size of the search space. With this the performance and the size of the solved problems were notoriously improved.

2 Reducing the Search Space by a Method Based in Ranges

The solving process in ASP consists of two stages: grounding and solving. It has been recognized that the grounding stage can represent a bottleneck because of the number of combinations frequently builds huge domains. For example, both in the planning and scheduling it is necessary to include the time. In this case, the resulting number of combinations becomes prohibitively large, especially when jobs require many processing units. Therefore, for planning and scheduling problems, the grounding stage is a highly demanding computational task and in the worst case, the task turns intractable.

A common problem solving methodology in ASP is to architect the program in three parts: GENERATE, DEFINE and TEST. The first part defines the collection of answer sets that represent potential solutions. Auxiliary concepts are defined in the DEFINE section. In our study case, this section is used to define the problem constraints. The TEST section consists of rules without head or constraints, used to eliminate answer sets that are not solutions.

After numerous intents to solve the AWS scheduling problem with the classic methodology used in ASP, it was evident that the problem could not be solved, being the largest possible one of 4 baths and 4 jobs. Looking for a more precise identification of the grounding trouble, the number of combinations experimentally obtained is N , and depends on the number N_b of baths, the number N_j of jobs, and the deadline D_d , according to the following formula.

$$N = D_d^{N_b * N_j} \quad (1)$$

For example, for a problem with 8 baths and 8 jobs and deadline equal to 120, the number of combinations is 1.1684e+133.

As is evident, the number of combinations is huge and we have to approach the problem in a different way. It is essential to find a way to cut down the grounding process. So, we decided to reduce the size of the set of possible solutions before implementing the generation rule. The application of this method requires to be extremely careful on the heuristics' selection that will be applied in order to reduce the number of combinations, because it is possible that sets of answers containing the optimal solution could be inadvertently deleted.

The selection of the heuristics is based on the knowledge of the problem and its solutions. In particular, the heuristic implemented is based on generating combinations of possible solutions through ranges. As the problem under study is about a flow workshop, all jobs must be processed sequentially and all jobs must start on the first resource. Therefore, it is useless to generate sequences with jobs starting its processing on the first bath later than half the time available to process the complete job because there would be insufficient time to finish the processing sequence. Since it is unknown in advance which jobs would be executed first, it was decided to do a formal modeling of the process. An advantage of this method is that the solution space is greatly reduced. The model consists on calculating ranges of sequences in which the solution is guaranteed to be found.

To do this, it is first necessary to calculate the sum of all processing times for each bath. This sum must be done for each job, according to the following equation:

$$\lambda = \sum_{k=1}^{nb} t, \forall j \quad (2)$$

where n_b is the number of baths, t is the processing time for each bath and j represents each job. Later, λ is used to find the maximum delay allowed for each job according to the next formula

$$\chi = \sum_{j=1}^{n_j} \text{deadline} - \lambda_j \quad (3)$$

where n_j is the number of jobs and λ_j is the total processing time for job j . Now, to calculate the range in which the combinations of each bath for a given job must be generated, it is necessary to know the total processing time from the first to the i -th bath, which is the previous bath whose range is being calculated according to

$$\rho = \sum_{k=1}^{th} t, \forall j \quad (4)$$

where th represents the i -th bath, and t is the processing time in each bath. Lastly, the superior and inferior limits of the range are calculated using the values previously obtained, as shown next

$$\Omega_{upper} = \chi_j + \rho_j, \forall j \forall k \quad (5)$$

$$\Omega_{lower} = \Omega_{upper_{j,k}} - \chi_j + 1, \forall j \forall k \quad (6)$$

3 Answer Set Programming Encoding

Equations 2 to 6 are encoded in the following rules:

```

dBsumOneJ(J, L) :- job(W), J = 1..W, L=#sum{D,B: bath(M),B=1..M,
    job(W), J = 1..W, duration(J,B,D)}.
maxDelayAnyJ(J,B,MaxR):- J=1..W, job(W), bath(M),B=1..M, dBsumOneJ(J,L),
    MaxR=deadline-L.
dSumJuntilB(J, Buntil, L) :- job(J), bath(Buntil),
    L=#sum{D,B: bath(Buntil), B=1..Buntil, duration(J,B,D)}.
range(J,B,Ru,Rl) :- bath(B), job(J), duration(J,B,D),
    maxDelayAnyJ(J,B,MaxR), dSumJuntilB(J,B,L), Ru=MaxR+L,
    Rl=Ru-MaxR+1.

```

The following is the generation rule, which calls the previous rules that calculates the ranges. In this way, a reduced version of the complete domain is obtained.

```

{do(J,B,T,D) :- times(L), T=1..L, T+D<=deadline, range(J,B,Ru,Rl),
    T >= Rl, T <= Ru }=1:- duration(J,B,D), job(W),
    J=1..W, bath(M), B=1..M.

```

Experimentally was found that applying the range method, the domain size is reduced to 98.5%. This percentage was calculated by first solving a π_1 program including only the generate rule. When π_1 program is solved, X number of models are found. Another π_2 program including rules related to equations 2 to 6 plus the generate rule was created. After the π_2 program is executed by the solver, Y number of models are found. Dividing Y by X and multiplying the result by 100 the referred percent is obtained. Given the huge size of combinations, the experiment was only tested on small problems.

The rules related to the resource uses were included In the **DEFINE** section. The rules named **assigned** are used to decide if it is possible to assign a job to some bath. A job can be assigned to bath B_2 if it was previously processed on B_1 , and B_2 comes immediately after B_1 , and processing between B_1 and B_2 is not overlaped or ending time is beyond of **deadline**. Rule **finished** is used to determine if a job has finished its processing in a certain bath.

```

assigned(J,B1,T1,D1):- do(J,B1,T1,D1), do(J,B2,T2,D2), job(J), T1<T2,
    D=T1+D1, B2==B1+1, T2>=D, D <=deadline.
assigned(J,B2,T2,D2):- finished(J,B1,T1), do(J,B2,T2,D2), B2==B1+1,
    T2 >= T1, F=T2+D2, F<=deadline.
assigned(J2,B,T2,D2):- finished(J,B,T1), do(J2,B,T2,D2),
    not finished(J2,B,T3),J !=J2, times(T3),
    T2 >= T1, F=T2+D2, T3 <=T2, F<=deadline.
finished(J1,B1,T2):- assigned(J1,B1,T1,D1), T2=T1+D1, T
    T1 < T2, T2 <= deadline.

```

We included a fourth section named **CONSTRAINTS** to the program. This section includes constraints such as: a) One of the scheduled jobs must start at bath one, rule **someJobStartAtOne**; b) baths must be processed sequentially, rule **anyJobStartSame**; c) policies ZW/NIS for chemical baths and LS for water baths, rules **delayedBw**, **neg.delayedBw**, **busyBath** and **badSincronyBch**; d) rules to detect overlapping in resources use also were included in this section, such as **overlapBathJobs**, **overlapJob-Baths** and **badStarting**; e) heuristic to avoid delay between a water bath and a chemical bath exceeding certain amount of time, rule **tooMuchDelay**, and; f) any job must finish after the deadline, rule **exceeds-Deadline**. The rules in question were defined as:

```

someJobStartAtOne(Mn):- Mn=#min{T1: times(T1), times(T2),
    do(J1,B,T1,D2), do(J2,B,T2,D1), B==1,T2 >T1}, Mn != 1.
anyJobStartSameT(T2):- do(J1,B,T1,D2), do(J2,B,T2,D1), J1
    J1 != J2, T2 == T1.

```

```

delayedBw(J2,B2,T2,D2):- finished(J2,B1,T1),
    assigned(J2,B2,T2,D2), B2==B1+1, B1\2 == 0, T2>T1,
    job(JT), times(E), not neg_delayedBw(J2,B2,T2,D2).
delayedBw(F):- finished(J1,B1,T1),
    F=#min{T2 : finished(J1,B1,T1),
    assigned(J2,B2,T2,D2), B2==B1+1, job(J2),
    times(T2), bath(B2) }, F !=#sup, F > T1.
neg_delayedBw(J2,B2,T2,D2):- finished(J2,B1,T1),
    assigned(J2,B2,T2,D2), B2==B1+1, B1\2 == 0,
    T2>T1, job(JT), times(E), not delayedBw(J2,B2,T2,D2).
neg_delayedBw(J2,B2,T2,D2):- busyBath(J2,B2,T2,D2).
busyBath(J2,B2,T2,D2):- finished(J2,B1,T1),
    assigned(J2,B2,T2,D2), job(JT), JT != J2,
    assigned(JT,B2,T3,D3), T2>T1, B1\2 == 0,
    B2==B1+1, F=#max{E: finished(J2,B1,T1),
    assigned(J2,B2,T2,D2), job(JT),
    JT != J2, assigned(JT,B2,T3,D3),
    E=T3+D3, finished(JT,B2,E)},
    T3 <=T1, F !=#inf, F==T2,F>T1.

badSincronyBch(T2):- do(J1,B1,T1,D1), do(J1,B2,T2,D2),
    B1\2 > 0, B2 ==B1+1, T2 > T1+D1.

overlapBathJobs(T2):- do(J1,B,T1,D1), do(J2,B,T2,D2),
    bath(B), duration(J1,B,D1), J1 < J2, T1 <= T2,
    T1 + D1 > T2.

overlapJobBaths(T2):- do(J1,B1,T1,D1), do(J1,B2,T2,D2), B2
    B2 == B1+1, T1<=T2, T1+D1 > T2.

badStarting(T2):- do(J,B2,T2,D2), do(J,B1,T1,D1),
    B2 > B1, T2 < T1+D1.

tooMuchDelay(J,B1,B2,F1,T2):- do(J,B1,T1,D1),
    do(J,B2,T2,D2), B1\2==0, B2==B1+1, F1=T1+D1,
    T2>F1, maxDur(M), L=T2-F1, L>2

exceedsDeadline(Tf):- do(J1,B1,T1,D1), do(J2,B2,T2,D2),
    T2 >=T1, Tf=T2+D2, Tf > deadline.

```

4 Experimental Results

The data used in this research were artificially created and published by Bhushan and Karimi [6]. This data has also been used by other researchers, such as [8] [2]. However, it was not possible use this data directly in ASP, because it contains real values and ASP only works with integers. Thus, the original data was multiplied by ten. After the problem was solved by ASP, the solution values were divided again by ten. The results shown in the next table were obtained through the experiment with and without the method based in ranges. In this experiment we used a small computer with an Intel processor Core i5.3337 CPU@1.80 Ghz X4, 6 Gb of memory, and 64 bits Ubuntu. The ASP solver used was clingo 4.4.

To test the hardware constraints on the solving process performance, the model based in ranges for problems of 6 baths with 6 jobs and 7 baths with 7 jobs was tested on a computer with CPU @2.88 Ghz, 8 cores, and memory of 16 Gigabytes. Only an optimal solution was found for the 6 baths - 6 jobs problem in a CPU time of 240.280 seconds and a makespan of 83. Neither the optimal solution nor the

Table 1: Results obtained with a typical rule and the range method proposed time is in seconds. NS: No solution

P:[BxJ]	makespan	CPU Time	Approach
[2x2]	18	1.820	No ranges
[2x2]	18	0.030	with ranges
[4x4]	50	96.920	No ranges
[4x4]	50	11.250	with ranges
[5x5]	–	NS	No ranges
[5x5]	65	165.890	with ranges

first solution was found.

5 Conclusions

ASP solvers limitation to solve big scheduling and planning problems comes from the approach used in the grounding step because a huge search space needs to be previously created for the reasoner to start looking for the solution. Although this problem is mitigated in some way by different strategies applied both in the grounding and solving stages, they are not effective enough in problems whose characteristics include variables like time, which turn the problem into one whose input grows exponentially. In our study case the product is processed sequentially making it convenient to exploit its structure to avoid generate useless combinations, like those starting at a certain time that would turn impossible to complete the processing. The proposed method only creates combinations with possibilities to finish on time and allowed to reduce the search space up to 98.5%. Although this reduction is important, it only allowed to increase the problem size from 4 bathrooms and 4 jobs to 6 bathrooms and 6 jobs. This results are not so surprising if we consider the example for a deadline of 83, with 6 jobs and 6 bathrooms, where the number of combinations was $8.0995e + 96$. Using the range method the problems size will still have $1.2149e + 95$ combinations.

Obviously, these kinds of problems need to be approached through a powerful methodology using a different strategy in the grounding step. It is foreseen that better results can be obtained approaching the problem with Constraint Answer Set Programming (CASP), because this hybrid methodology uses a completely different strategy in the grounding step. However, this approach is motive of another research what would be published in the future. Nonetheless, we consider that the proposed method is applicable to small flowshop scheduling problems as well as to establish a way to exploit the problem structure in order to reduce the number of possible solutions and improve the solver performance.

References

- [1] A. Allahverdi (2016): *A survey of scheduling problems with no-wait in process*. *European Journal of Operational Research* 255(3), pp. 665–686, doi:10.1016/j.ejor.2016.05.036. Available at <https://www.sciencedirect.com/science/article/abs/pii/S0377221716303733>. Note.
- [2] P.M. Castro A.M.Aguirre C.A. Mendez (2011): *A novel optimization method to automated wet-etch station scheduling in semiconductor manufacturing systems*. *Computers and Chemical Engineering* 28(Num-

- ber):4, doi:10.1016/S1570-7946(10)28148-8. Available at <https://www.sciencedirect.com/science/article/pii/S1570794610281488>. Note.
- [3] R. Uzsoy C.D. Geiger K.G. Kempf (1997): *A Tabu search approach to scheduling an automated wet etch station*. *Journal of Manufacturing Systems* 2(Number):4, doi:10.1016/S0278-6125(97)85674-9. Available at <https://www.sciencedirect.com/science/article/pii/S0278612597856749>. Note.
- [4] N. Leone F. Ricca F. G. Grasso M. Alviano-M. Manna V. Lio S. Liritano (2012): *Team-building with answer set programming in the gioia-tauro seaport*. *Theory and Practice of Logic Programming* 12(3), pp. 361–381, doi:10.1017/S147106841100007X. Available at <https://arxiv.org/abs/1101.4554>. Note.
- [5] J.M. Novas G.P. Henning (2010): *Reactive scheduling framework based on domain knowledge and constraint programming*. *Computers and Chemical Engineering* 34(12):4, doi:10.1016/j.compchemeng.2010.07.011. Available at <https://www.sciencedirect.com/science/article/pii/S0098135410002590>. Note.
- [6] S. Bhushan I.A. Karimi (2003): *An MILP approach to automated wet-etch scheduling*. *Industrial and Engineering Chemistry Research* 42(7):4, doi:10.1021/ie020296c. Available at <https://pubs.acs.org/doi/abs/10.1021/ie020296c>. Note.
- [7] S. Bhushan I.A. Karimi (2004): *Heuristic algorithms for scheduling an automated wet-etch station*. *Computers and Chemical Engineering* 28(3):4, doi:10.1016/S0098-1354(03)00192-3. Available at <https://www.sciencedirect.com/science/article/pii/S0098135403001923>. Note.
- [8] C.A. Mendez L.J. Zeballos P.M. Castro (2011): *Integrated constraint programming scheduling approach for automated wet-etch stations in semiconductor manufacturing*. 20. *Industrial and Engineering Chemistry Research* 50, pp. 1705–1715, doi:10.1021/ie1016199. Available at <https://pubs.acs.org/doi/abs/10.1021/ie1016199>. Note.
- [9] M. Barry M. Nogueira M. Balduccini M. Gelfond R. Watson (2001): *A prolog decision support system for the space shuttle*. *Proceedings of the 3rd International Symposium on Practical Aspects of Declarative Languages 1990 of LNCS*(Number):4, doi:10.1007/3-540-45241-9_12. Available at <https://dl.acm.org/citation.cfm?id=667928>. Note.
- [10] M. Runge M. Gebser P. Obermeier T. Schaub M. Ratsch-Heitmann (2018): *Routing Driverless Transport Vehicles in Car Assembly with Answer Set Programming*. *Theory and Practice of Logic Programming* 18(3):4, doi:10.1017/S1471068418000182. Available at <https://www.cambridge.org/core/journals/theory-and-practice-of-logic-programming/article/routing-driverless-transport-vehicles-in-car-assembly-with-answer-set-programming/6965B3B4C0D5F5BB0D4BB344582954A4>. Note.
- [11] M. Gelfond V. Lifschitz (1988): *The stable model semantics for logic programming*. *Fifth International Conference on Logic Programming*:4. Available at https://www.researchgate.net/publication/2408055_The_Stable_Model_Semantics_For_Logic_Programming.