Cruzando Fronteras:
Tendencias de Contabilidad Directiva para el Siglo XXI

DETERMINISTIC CROWDING IN GENETIC ALGORITHM TO SOLVE A REAL-SCHEDULING PROBLEM
PART 2: APPLICATION

Mª Elena Pérez-Vázquez
Angel M. Gento-Municio
Organización y Gestión de Empresas.
E.T.S. Ingenieros Industriales.
Universidad de Valladolid. Spain.
elena@eis.uva.es - gento@eis.uva.es.

Resumen

The genetic algorithms (GAs) belong to the metaheuristics approaches, that is, they need to be adapted to the characteristics of the problem. In this work we will develop the realized implementation to solve a real and complex scheduling problem. Due to the high complexity, the problem has been mainly divided in three independent parts. For them, we will need to establish the parameters system of the GAs that direct the search. To prove the power and utility of the developed process we will analyze two practical instances that the enterprise had to solve by a previous method and the results obtained will be compared.

Palabras clave: Algoritmos genéticos, problema real, programación de operaciones.

ACODI

VII Congreso del Instituto Internacional de Costos

UNIVERSIDAD DE LEÓN

II Congreso de la Asociación Española de Contabilidad Directiva

1.- Our Implementation

Due to the high difficulty of the production process (see Part 1) we have divided the problem in three phases:

Phase 1: In this first phase we realice the study of the unsettled orders and their due date. With these facts we will obtain the production scheduling within a week (or month)'s time (period of time considered by the expert). Also, we will try to balance the production of the following weeks (or months). Therefore, it is a phase that assigns the production orders (sidings and sleepers to manufacture) and balances the charges.

Phase 2: With the sleepers obtained in the phase 1, we must study their necessary sequence to achieve the objective of maximizing the usefulness of the installations measured in produced lineal metres and producing all of the orders.

The purpose of this stage is to optimize the sequence of sleepers on the manufacturing lines (plank moulds), considering to these plank moulds as a succession of lines, just as it is shown in the Figure 1.

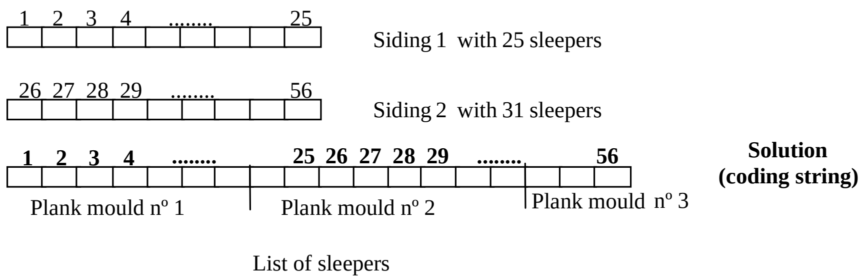


Figure 1: Example of coding process for two sidings to be manufactured

In this way, the coding process of the GA uses the permutation of integer and consecutive numbers (permutational coding), where each number will correspond to a certain sleeper. The total size of the coding string will be the sum of the units of each sleeper selected in the phase 1 to be manufactured (Figure 1). Each possible permutation of all these integer numbers is a solution of the search space that the GA will modify to maximice the objective of this phase.

The initial population will be generated in a random way. Initially, *N* random permutations of the integer numbers will be created. *N* being the size population (one of the parameter of the GA that must be fixed) we will use a value of 50 individuals.

To calculate the quality (fitness function), first of all, we must make sure the solution is feasible, that is, we must check if the solution represented by the coding string does not violate the constraint of the maximum number of existing moulds for each sleeper. If this constraint is not accomplished a new solution must be created.

We take account that each manufacturing lines measures 104 metres. For which, we divide the solution (list of sleepers) in sub-list of sleepers whose length sum less than these 104 metres (we note the length of separator). Thus, we will know how many plank moulds we will need to manufacture (Figure 1). Then, within each plank mould we check if the constraint of mould is accomplished (perhaps the solution includes two units of the same sleeper in one plank mould but we only have one mould of it), we can not use more moulds that we have. If the solution accomplishes this constraint is called feasible, if not, we will have to generate other.

Then, we calculate the percentage of occupation of each plank mould and the total average percentage; first objective to maximize in this phase.

Phase 3: This phase deals with the better sequences of sleepers obtained by the GA defined in phase 2. Each one of these solutions constitutes the starting point of the current phase, in such a way that the process will go on with as many genetic algorithms as different solutions suggested in phase 2. Not only with the best solutions, but with all different solutions, since there is an inverse relation between this objective and the two ones of the next phase.

Thus, we can offer to the expert different final solutions with fitness very similar for what he can evaluate them taking account some constraints as breakdown machine or others non-quantifiable objectives.

The coding process is similar that the previous algorithm excepting that in this case, each integer number represents a plank mould, that is, the set of sleepers that forms a manufacturing line (see Figure 1). The initial population of the algorithm *i* is formed by the not repeated solution *i* of the before phase and its possible permutations, thus the size of the population is completed (50 individuals, the original and its 49 random permutations).

To calculate the fitness, we will distribute the groups of plank moulds in the different existing lines. For this, we need know how many manufacturing lines are going to be disposed. Of the four manufacturing lines that we have, we will use the four, three, two or only one of them in function of the manufacturing conditions. To distingue the different concepts we have called series to the manufacturing lines that we can work on a time (Figure 2).

Thus, we can calculate the number of necessary moulds of each type for each series. If this number is higher than the existing moulds the solutions will be not feasible and it will be necessary to create other, exchanging the solution until a valid one is obtained. Afterwards, we will determinate the change of moulds required from one series to the next. The final fitness will be the percentage of common existing moulds between groups.

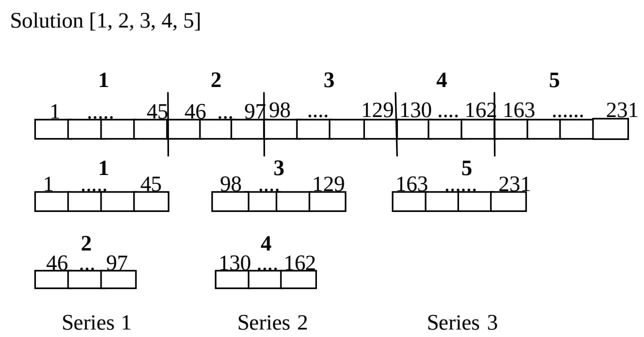


Figure 2: Solution divided in series with two manufacturing lines

On the other hand, there is a second objective that we must take account: the number of manufacturing lines, which are actually used. When the expert indicates us that we can use a certain number of manufacturing lines for the limitations of the labour or some technical problem, it is a maximum number. Due to the constraint of the number of existing moulds, it is possible that we can only fill up some of this maximum number and therefore, the percentage of utilization will not be of the 100%. For this case, we have used a second objective, the percentage of number of occupied manufacturing lines. This value and the percentage of repeated moulds between series are sum, changing the problem of two objectives into a problem of one objective.

The final result of the process of both phases is a solution that maximizes the utilization of the manufacturing resources and minimizes the number of changes.

To finish this section, the process general of the two genetic algorithms used to solving this real problem is shown in the Figure 4.

We want to note that the generation number and size population are parameters that must be adapted to the problem complexity. In our case, taking account the characteristic of the instances (section 2) we have used 10.000 and 7000 generations for the first and second instances. The population size has been fixed to 50 solutions. And, we have used a 80% and 20% percentage for the crossover and mutation operators, respectively.

2.- Description of the Instances

The first instance that we will study is a real scheduling problem of this enterprise. It is composed of the sleepers that must be fabricated for a certain month (May, 2000).

This problem is composed for the three different configurations of sleepers with the following characteristics (see Tables 1 and 2). The dark shaded rows indicate the sleepers that have a constraint of moulds:

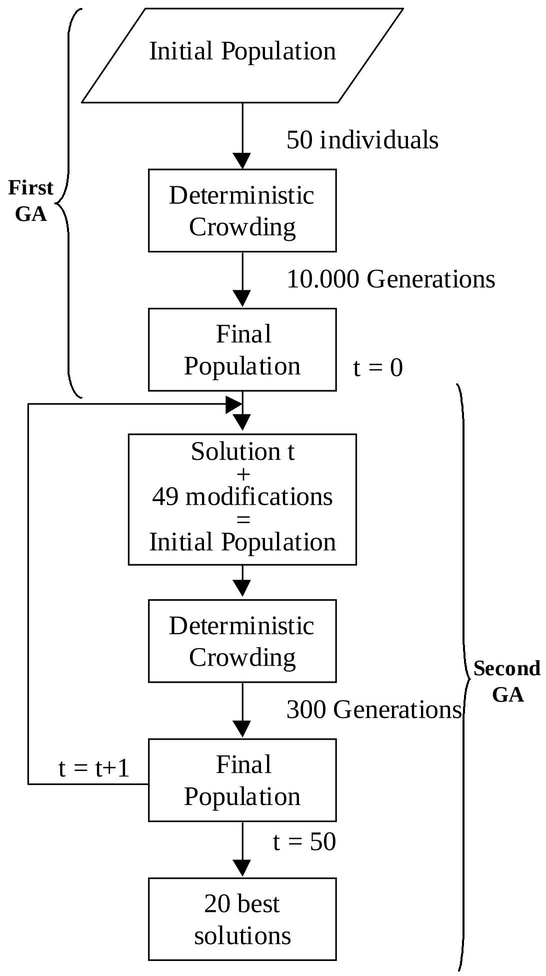


Figure 4: Global process of the two concatenated GAs

Type (Configuration)	**1**	**2**	**3**
Metres	286.42	1554.18	913.60
Sleepers Number	89	461	352

Table 1: Metres and number of sleepers of each configuration

In the Table 2, the first column shows the number of different sleepers with the characteristics: number of available moulds (constraint) and number of necessary units of each one of these sleepers. For example, the first row shows us that we must fabricate 2 units of each one of these 12 different sleepers which have only 1 available mould of them, that is, there exist a constraint to manufacture these 12 sleepers.

	Sleepers	**Number of available moulds**	**Units of each sleeper**
Type 1	12	1	2
	18	1	3
	1	5	2
	1	2	4
	1	5	5
Type 2	96	1	96
	103	1	206
	3	1	9
	31	1	124
	2	2	4
	1	10	20
	1	3	1
	1	2	1
Type 3	1	5	1
	1	78	351

Table 2: Characteristic of the first instance

The second instance is a very specific problem that the enterprise has to face many times. In this case, the sleepers are only of the type 3, it is only for the AVE (see Tables 3 and 4).

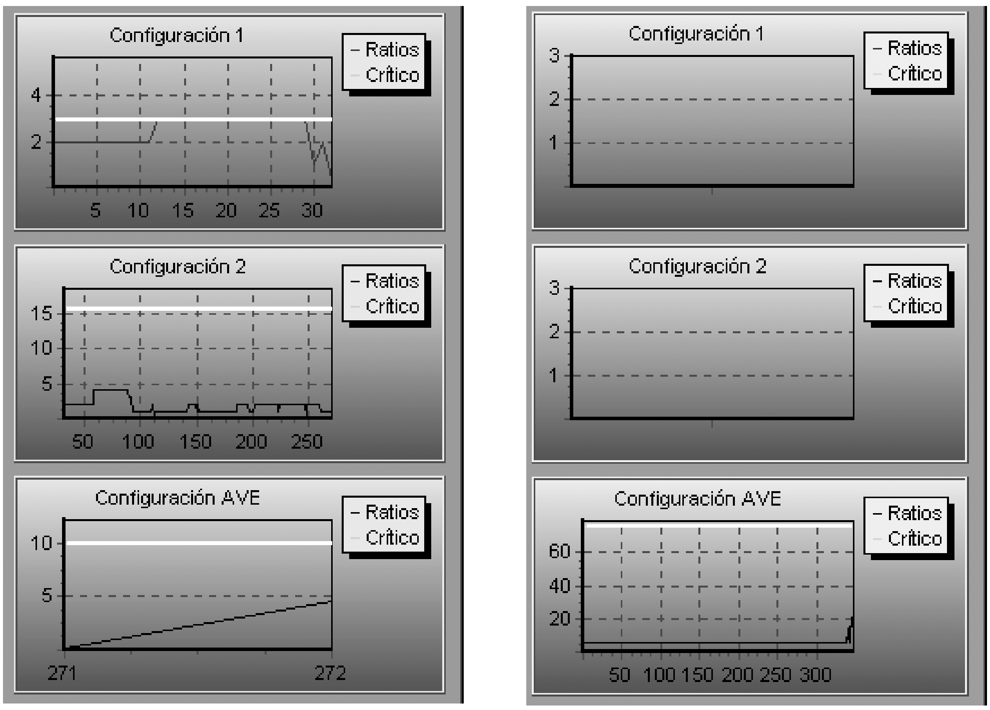
Type (Configuration)	**3**
Metres	7591.19
Sleepers Number	3095

Table 3: Metres and number of sleepers for the AVE problem

The program allows doing a previous study of the limitations of moulds. In the Figure 5 is shown the graphic relation between the necessary number of plank moulds to manufacture according the number of existing sleepers (white line), with the necessary number of plank moulds (critical) in function of the number of available moulds (blank line). If we must realize two units of the same sleeper and we only have one mould, we will need allocate the two units in two different plank moulds.

	Sleepers	Number of available moulds	Number of necessary sleepers
Type 3	1	1	20
	338	1	5
	1	2	10
	1	2	20
	1	4	85
	1	5	50
	1	5	70
	1	6	55
	1	15	165
	1	17	125
	1	78	805

Table 4: Characteristic of the AVE problem



(a) (b)

Figure 5: Study of critical number of plank moulds

As we can see, the one type of the first instance (Figure 5a) the critical number is the same for some sleepers than the necessary number of plank moulds, for which this problem has a high difficulty. For the second instance (Figure 5b), the critical number is higher than the necessary number for that the problem is difficult for the size but not for the constraints.

3.- Analysis of Instances and Results

3.1.- First Instance

The real dates for this instance offered for the enterprise are summarized in next table. As we can see, for the one and two configurations the enterprise did not use ever the four manufacturing lines at a time because the old scheduling system can not control all lines for the limitation of the moulds and the complexity of the problem. For this, they used three series with only one manufacturing line, and ten series with two manufacturing lines, that is, 13 series and 23 plank moulds. For the AVE, the real needs are 351 units of the same sleeper of which we have 78 moulds, for which there are the necessary moulds to fill at a time only two manufacturing lines.

Dates	**Manufacturing lines for 1 and 2 types**				**Manufacturing lines for AVE**			
	L1	**L2**	**L3**	**L4**	**L1**	**L2**	**L3**	**L4**
25-4-2000					X	X		
27-4-2000					X	X		
28-4-2000				X				
2-5-2000				X				
3-5-2000					X	X		
4-5-2000			X	X				
5-5-2000	X	X						
8-5-2000			X	X				
9-5-2000	X	X						
11-5-2000			X	X				
12-5-2000	X	X						
13-5-2000				X				
15-5-2000			X	X	X	X		
17-5-2000	X	X						
18-5-2000			X	X				
19-5-2000		X		X		X		

Table 5: Real dates of the first instance

For the AVE configuration, the solution is the same than the obtained by the GA. The convergence was obtained in the 3 or 4 generation (for the 12 runs performed). However, due to the constraint that affects us in this case (available moulds) the obtained results for the one and two configurations by the deterministic crowding GA are very different.

For the second phase, where we fill up the plank moulds with the objective of maximizing their utilization. The average best percentage obtained for 12 replicas is shown in the Table 6.

Number Generations	**Type 1**		**Type 2**		**Type 3**	
	Av.	**Var.**	**Av.**	**Var.**	**Av.**	**Var.**
1-100	0.5458	0.0002	0.9406	0.0001	0.99	0.000
101-200	0.6031	0.0004	0.9615	0.0000	0.99	0.000
201-400	0.6681	0.0005	0.9762	0.0000	0.99	0.000
401-600	0.7219	0.0015	0.9815	0.0000	0.99	0.000
601-800	0.7617	0.0019	0.9841	0.0000	0.99	0.000
801-1000	0.7960	0.0020	0.9854	0.0000	**0.99**	0.000
1001-1200	0.8176	0.0015	0.9864	0.0000	0.99	0.000
1201-1400	0.8345	0.0014	0.9872	0.0000	0.99	0.000
1401-1700	0.8552	0.0015	0.9880	0.0000	0.99	0.000
1701-2000	0.8698	0.0018	0.9886	0.0000	0.99	0.000
2001-3000	0.8954	0.0022	0.9897	0.0000	**0.99**	0.000
3001-4000	0.9087	0.0026	0.9904	0.0000	0.99	0.000
4001-5000	0.9155	0.0028	0.9909	0.0000	**0.99**	0.000
5001-6000	0.9190	0.0029	0.9912	0.0000	0.99	0.000
6001-7000	0.9218	0.0030	0.9915	0.0000	**0.99**	0.000
7001-8000	0.9242	0.0031	0.9917	0.0000	0.99	0.000
8001-9000	0.9263	0.0032	0.9919	0.0000	0.99	0.000
9001-10000	0.9284	0.0033	0.9921	0.0000	**0.99**	0.000

Table 6: Average best percentages of filling for each type of configuration (12 replicas)

Each solution obtained by the second phase after 1000, 3000, 5000, 7000 and 10000 generations will be used as start point of the third phase, where the plank moulds will be allocated in the series. For this, we will consider the best 20 solutions of each one of the 12 performed replicas for the second phase.

We can not only take account the solutions of the last generation (10000) because we do not know how the filling percentage affects to the placing in the series. At first sight, the relation will be inverse, if the filling of the plank moulds is higher, then the allocation could be harder than with others less filling plank moulds. Perhaps, with worse solutions of the second phase, we can obtain best results in the third phase.

The results for the 1 and 2 types are summarized in table 7.

Generation Number	Average Percentage	Average moulds	Group 1	Group 2	Group 3	Group 4	Series
1000	46.35	74.73	3.54	5.28	1.75	0.16	10.73
3000	45.51	77.01	3.80	5.14	1.60	0.17	10.71
5000	45.32	77.79	3.87	5.07	1.60	0.17	10.71
7000	44.99	74.80	4.06	4.93	1.58	0.20	10.77
10000	44.70	74.33	3.93	5.30	1.47	0.13	10.82

Table 7: Average results of all solutions (20 solutions for 12 replicas), first instance

The column of average percentage shows the average occupation of the used series. The column of average moulds shows the average number of common moulds between series.

The groups show the number of series with: only one occupied manufacturing line (group 1); two occupied manufacturing line (group 2); three (group 3); and four, all occupied manufacturing lines (group 4).

The column of series shows us the number of different necessary series to fabricate all plank moulds obtained in second phase.

Just as we have commented before, if the percentage of utilization within each plank mould is higher, then the difficulty to fill the four available manufacturing lines in each series will be higher. Thus, with the obtained solutions after the 5000 generation the results are better than with the last solutions (10.000 generations). The number of manufacturing lines that are used within each series is lesser and therefore, the number of common moulds also decreases.

However, as the second GA is directed by two objectives, it seems more reasonable to consider only the non-dominated of the 20 best solutions. In this case, the results are shown in the Table 8.

Generation Number	Average Percentage	Average moulds	Group 1	Group 2	Group 3	Group 4	Series
1000	49.39	85.97	2.76	5.07	2.01	0.26	10.10
3000	47.66	89.94	3.16	5.02	2.04	0.09	10.30
5000	47.98	87.88	2.81	5.47	1.75	0.16	10.16
7000	47.16	85.61	3.64	4.49	2.00	0.24	10.37
10000	46.57	84.47	2.98	5.84	1.67	0.00	10.43

Table 8: Average results of non-dominated solutions (12 replicas), first instance

As we can see, if only we take account the non-dominated solution, the average quality will be higher, and the number of series will be lesser.

For example, in the table 9 we show one solution with the quality similar to the average in the 5000 generation of the first genetic AG and 300 generation in the second GA (Average), other with the lesser number of series and plank mould used (Best), and others two with only two manufacturing lines used as in the real dates (Two lines):

Solution	Average Percentage	Average moulds	Group 1	Group 2	Group 3	Group 4	Series	Plank mould
Average	47	93	5	1	4	0	10	19
Best	52	108	2	4	3	0	9	19
Two lines	50	65	0	10	0	0	10	20
	47	104	1	9	0	0	10	19

Table 9: Results of a representative solution.

As we can see, the lesser number of necessary days to manufacture all sleepers is 9, against the 13 that the enterprise needed.

As the information of the "Colegio Oficial de Aparejadores y Arquitectos Técnicos" the salary of the different works in the plant is:

Category	**Number of persons**	**Salary**
Specialized unskilled	2	1824 pts/h.
Assistant of plant	1	2011 pts/h.
First Officer	1	2145 pts/h.

Table 10: Salary of the works

In this way, the cost of the daily salary is 62.432 pts. Therefore, the results offered for the GA allow us a saving of salary of the 249.728 pts.

However, this is not the only saving possible. We must take account that the decrease of the number of realized plank moulds is associated to the decrease of the number of used steel bar. In our case, the steel has a diameter of 55 mm and a cost per lineal metre of 126,26 pts. As in each plank mould we use 9 or 12 bar of the 104 metres, the real fixed cost in steel bar is the 2.265.104,4 pts. However, with our solution the fixed cost will be 1.995.918,08. That is, the saving is the 269.186,32 pts. Due to which the total saving of salary and steel bar is of the 518.914,32 pts.

3.2.- Second Instance

In this case, the enterprise offers us a problem very common in the manufacturing process. It consists in five units of the same siding.

The convergence of the evolution in this instance is very fast (Figure 11). After 1000 generations the filling percentage is of the 90.4% and the increase after 6000 generations is the 0.97%. For this, the fixed generation number is less than for the first instance.

Number Generations	**Type 3**	
	Av.	**Var.**
1-1000	0.9042	0.0738
1001-2000	0.9087	0.0745
2001-3000	0.9106	0.0748
3001-4000	0.9051	0.0812
4001-5000	0.9122	0.0751
5001-6000	0.9126	0.0752
6001-7000	0.9130	0.0752

Table 11: Filling percentage for the second instance

The need of the enterprise about this problem is to obtain less than 78 plank moulds. In this case, the second optimization (allocating these plank moulds in series) is not necessary, and so, we have only to run the first and second phases, and only the first GA. The number of solutions, which number of plank moulds is 77, 76 and 75 for each 1000 generations is shown in the table 12.

	Generations						
Plank Moulds	**1000**	**2000**	**3000**	**4000**	**5000**	**6000**	**7000**
75	11.50	15.58	26.41	33.41	37.33	40.50	42.08
76	36.50	33.91	23.50	16.58	12.66	10.36	8.63
77	8.45	1.50	0.50	0.00	0.00	0.00	0.00
Av.	75.94	75.72	75.48	75.33	75.25	75.20	75.17

Table 12: Average number of necessary plank moulds (12 replicas)

We can see that after 1000 generations the need of the enterprise is yet accomplished. All of the solutions in the population have a number of plank moulds less than 78, being very important the evolution of the number of solutions with 75 plank moulds. These solutions have never obtained for the enterprise with the old scheduling system.

In last row, the average number of necessary plank moulds for the 50 solutions of the 1000, 2000, ..., 7000 generations are shown. We see as this number decreases according the process evolves.

The main cost saving is refereed to the number of steel bar used in the fabrications of the plank moulds. The salary saving is not known since we do not obtain how many series are necessary to do for this solution.

Concluding Remarks

The first question that always put us when we must solve (optimize) a combinatorial problem is what the best optimization method is. The large development of different metaheuristic process makes difficulty the best selection. For this, the question is not what the best is, but what the best adapted to the problem is.

For this real and complex problem, where there are different constraints and multiple objectives, we think the best adapted are the GAs. And the obtained results support this decision.

It has been proved that the GAs adapt to the handle of different constraints. They are very flexible to incorporate new constraints in the future without main changes, and the process permits to introduce new resources (moulds) and new products (sidings).

Likewise, the results obtained improve the results that the enterprise obtained with the old scheduling system. This improvement arises a cost saving of salary and utilized resources that allow increasing the profit of the enterprise and accomplishing the global strategy of rise and quality.