Production Scheduling on Practical Problems

Marcius Fabius Henriques de Carvalho and Rosana Beatriz Baptista Haddad
Pontifícia Universidade Católica de Campinas–PUC-Campinas, Centro de Tecnologia da Informação Renato Archer-CTI, Brazil

1. Introduction

Management’s desire to be more competitive and to increase profits through manufacturing is evident. Customer responsiveness, increased output, lower manufacturing costs, better quality, short cycle times, bottleneck control and operational predictability, among many other themes, are hot issues on manager’s minds. The management of manufacturing processes is a complex problem that’s objective is to sell goods and services to the marketplace, through internal production resources and supplier agreements and capabilities. It is therefore advisable to structure the solution of the problem hierarchically, considering different aggregation levels of information and decisions, Figure 1.

Fig. 1. Production planning management hierarchy

Effective enterprise management decisions at any level must be supported by modern tools and techniques. Both hierarchy and tools can offer tremendous potential advantages when adequately integrated into the search for a good solution. In long-term operation planning the strategic business issue, generally looking five or more years ahead, aims to provide for the long-term capacity requirements and resource allocations necessary to meet the future organizational objectives. This is done by planning for capacity changes in line to meet plans for new products/services, technologies and markets (Hill, 1991). Strategic decisions search for two kinds of production flexibility, flexible technology and flexible capacity, to meet long term demand fluctuations. Each strategy is underpinned by a set of operations decisions on technology level, capacity amount, production quantity, and pricing strategy evaluation determines how market uncertainty, production cost structure, operations
timing, and investment costing environment affect a firm’s strategic decisions (Yang, Cheng, 2011).

At this level, products with similar production costs and seasonality are grouped together and considered as one family. This procedure subsidizes purchasing, resource development, production and customer service policies and supports the generation of a long-term production plan to guide broad courses of actions that an organization needs to follow in order to achieve its objectives, (Bitran et al 1981,1982; Ozdamar, et al 1998).

The medium-term or aggregated planning covers a period of up to two years. The technical feasibility of implementing the strategic level decisions, taking into account time and available resources, are verified. Production plans for product families sharing the same setup costs, that reduce the final production cost, are generated. At this level the data are less aggregated than they are at the strategic level. This makes it possible to generate a plan with a gross level of overtime or subcontract production capacity for the medium-term horizon considered, for example, a few weeks, (Caravilla & Souza, 1995). At this level, the production plan extends beyond the enterprise resources to consider downstream activities as distribution, sales and inventory based on customer market information and upstream activities as supplier contract and constraints, while observing all relevant internal constraints. The production plan connects sources and sinks through material flow equations and is usually driven by economic objectives (Mutairi, 2008).

In the short term the production planning data is broken down into tasks to be performed in the short term, for example, a day or a few days. Detailed demand and processes are considered. At this level, sequences are decided on the factory floor, which generally consider several production resources responsible for processing items at the various production stages allocated to each different product (Karimi, B. et al. 2003; Carvalho & Silva Filho 1996). On a fictitious factory floor, dedicated to the production of only one item, with infinite availability of inputs and production resources, it should be relatively simple to coordinate the supply of raw material with the production of components to meet the demand requirement. However, in practice this situation does not occur. Real systems are complex environments where various items compete for limited capacity of the available resources.

Production scheduling deals with operational decisions at each plant, such as the machine production sequence. Its purpose is to transform the short term production plan, utilizing the available resources over a given time horizon into a schedule useful for all operations within a time horizon of a few days. A good production schedule is one that has the ability to synchronize the job chain network, starting on projected and updated completion dates.

The scheduling problems in general can be modeled using continuous or discrete time models. This chapter discusses the discrete time models. They are based on two features. First, the scheduling horizon is divided into a finite number of time intervals with predefined duration. Second, the events are allowed to happen only at the boundaries of the discrete intervals. Hence, the complexity of the problem is reduced and the solution becomes less cumbersome.

Production scheduling methods can be further classified as static scheduling and dynamic scheduling. Static operation scheduling is performed during the compilation of the application. Once an acceptable scheduling solution is found, it is deployed as part of the application image. In dynamic scheduling, a dedicated system component makes scheduling decisions on-the-fly (Wang, Gong, Kastner, 2008).
This chapter is specifically related to the discrete time and static production scheduling problems. Scheduling is typically driven by feasibility and focuses on short-term time horizon. The goal of scheduling is to orchestrate an optimized behavior of a resources-limited system over time, considering the predictable, uncertain and dynamic variables of a real system. Although the research community has considered the scheduling as a solved problem, such a conclusion presumes a rather narrow and specialized interpretation of the problem and ignores much of the process and context in practical environments. The design of more tightly integrated planning and scheduling in real systems is a question that still requires research to reach a practical solution. This theme is considered in this text through the analysis of two auto-part industry applications.

2. Model definition

The real world can be represented by models in order to better understand and propose a solution for some of its particular problems. The models should express the decision maker’s needs and present results in a way that they can be easily read and understood (Voβ and Woodruff, 2006). Also, to find a solution to the problem, it is necessary to consider the level of detail to be included in the model, and the computational effort that must be made to solve it.

For example, by definition, we always prefer a better value of the objective function. However, having the best may not be so important if its cost is high and another solution that satisfies the constraints is enough for the purpose of the decision. In another case, the objective can be of considerable importance and the best solution should be pursued. In either case, we must find a balance between the difficulty in obtaining the result and the desired objective.

In order to define the model, we have to analyse the problems and characterize them according to a variety of aspects and classification criteria. One aspect is the consideration of an integer variable (Wu at. al. 2011) or not since the greater the number of integer variable the harder it will be to solve the problem. But, in some cases it is sensible to require some variables to take on the integer value (Voβ and Woodruff, 2006).

For example, if the product is of a high value, with low production intensity, it does not make sense to talk about a partial product. But if the product is of a low value and higher production intensity, as in the auto part production industry, the integer consideration is not necessary. Also, if the setup time is very high it is necessary to work with integer variables. If however the enterprise is in the middle of a supply chain and receives large weekly orders it may be preferable to process the entire order rather than splitting it into a small one.

Nothing that has anything to do with planning is truly linear, though non-linearity in the variables can make the solutions harder to obtain. But, to integrate the planning model with scheduling a linear model is enough in order to obtain the necessary solution (Voβ and Woodruff, 2006).

Infinite scheduling allows more than one job to be scheduled for production within a limited capacity resource at the same time. When many jobs are scheduled at the same time, which is often the case, the same limited resource is in a state of scheduled chaos. The only alternative is a classic scenario: supervisors and production control personnel must take hands-on control of the shop floor scheduling.

A Finite schedule is one that makes sure that simultaneous schedule contention for the same capacity resource is avoided. Starting from a non-capacitated schedule, a temporal capacitated schedule can be met by rescheduling, following some heuristics, Figure 2.
One way of rescheduling is anticipating or postponing a production order according to priority rules. For example, on a shop floor that produces for both make-to-stock and make-to-order, a “make to stock” job will have lower priority than a “make to fill an order” job. However, a question arises in a system where a temporal solution is needed for multiple customer orders of differing sales values, simultaneously competing for multiple resources. How can an order in this complex environment be prioritized?

Also, distinction can refer to deterministic versus stochastic models. While in deterministic models all data are known in advance, in stochastic models data are based on distributions or on a measure of uncertainty (Voß and Woodruff, 2006). Static models assume that parameter values do not change over the planning horizon (e.g. a continuous demand at the same rate in every period) while dynamic models allow for variations. The planning horizon can be assumed to be finite or infinite. Costs to be considered are holding costs, setup costs, or production costs.

Production scheduling models can consider one or multiple products. The multi-product consideration implies the development of a multiproduct production plan in a scarce resource environment. The problem can also be single stage or multiple stage and all the above problems can be represented by a capacitated model that recognizes that the capacity of each resource is defined by a finite number or amount. In the situation where there is not enough capacity, certain decisions have to be made, such as those regarding backordering or loss of sale policy.

In any of the above situations, the development of a model must take into account the following points:

- The model must be simple;
- It is advisable to avoid the development of mega-models;
- It is advisable to use similarities and analogies

In the practical application, one can either apply an approximate method that delivers a good solution in an acceptable time or an optimization procedure that yields a globally optimal solution, solving both timing and resource constrained scheduling problems, but
requiring a very high computing time (Jain and Meeran, 1999). Another option can even be to integrate both procedures, taking specialized heuristic to treat selected constraints and objectives in an ad hoc manner.

Generally speaking, this would mean taking advantage of a problem-specific engineering solution to obtain a result that meets a given application’s requirements. There continues to be great need for research into techniques that operate with more realistic problem assumptions (Smith, 2003). To exemplify the above concepts the following section discusses a number of modeling aspects in more detail.

3. The scheduling problem

Production scheduling can be defined as the allocation of available production resources over time to best satisfy some set of criteria. Scheduling problems arise whenever a common set of resources (labour, material and equipment) must be used to make a variety of different products during the same period of time (Rodammer, White, 1988). The objective of scheduling is to efficiently allocate resources over time to manufacture goods such that the production constraints are satisfied and the production costs are minimized. This involves a set of tasks to be performed, and the criteria may involve both tradeoffs between early and late completion of a task, and between holding inventory for the task and frequent production changeovers (Graves, 1981).

Operations require machines and material resources and must be performed according to a feasible technological sequence. Schedules are influenced by diverse factors such as job priorities, due-date requirements, release dates, cost restrictions, production levels, lot-size restrictions, machine availabilities, machine capabilities, operation precedence, resource requirements, and resource availabilities.

To illustrate a production schedule, consider a one-product, three-machine job-shop scheduling problem, shown in Figure 3. The simplest scheduling occurs when we have unlimited capacity resources for the given application while trying to minimize its latency.

Fig. 3. One-product multi-stage production system
For this problem, simply solve it is to schedule an operation as soon as all of its predecessors tasks have completed, which gives it the name As Soon As Possible. It is closely related with finding the longest path between an operation and the demand due date. The first observation is that the maximum capacity of this system is equal to the lesser of the maximum capacities of the resources.

Now consider Figure 4, with the same three machine system, processing three different products under distinct production sequences. In this case, the maximum productive capacity of the system depends on the mix and sequence of production, the rate of resource sharing and the time profile of demand. Therefore, the production capacity of any multiproduct, multistage system is dependent on the mix of the demand and on the production schedule.

![Multi-stage multi-period multi-product production system](image)

This problem grows in complexity when considering various time periods for manufacturing resource planning. Note that the required ordering of operations within each job (the sequence technology) is preserved and the ordering of operations on the machine was selected to minimize the total production time.

To illustrate the problem of resource allocation, Table 1 provides routes data and the processing time of a cell production, composed of three machines. The Prod1 first operation occurs on the machine M3, and requires one unit of processing time. The second operation is on the machine M1, requiring seven units of time. Finally, the third operation requires 6 units of time and occurs on machine M2. The production sequence for products 2 and 3 follow the same logic.

<table>
<thead>
<tr>
<th>Operation Sequency</th>
<th>Prod1</th>
<th>Prod2</th>
<th>Prod3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/M3</td>
<td>6/M1</td>
<td>3/M3</td>
</tr>
<tr>
<td>2</td>
<td>7/M1</td>
<td>5/M3</td>
<td>6/M2</td>
</tr>
<tr>
<td>3</td>
<td>6/M2</td>
<td>3/M2</td>
<td>10/M1</td>
</tr>
</tbody>
</table>

Table 1. Three product production sequence
One of the Gantt charts for production allocation of the three products is shown in Figure 5. It is assumed that all items are available at the beginning of the process and that operations are not shared.

![Gantt chart for three products production allocation](image)

**Fig. 5. Gantt chart for three products production allocation**

However, when considering a real problem it is not enough to simply have a heuristic to allocate the production of the items to their machines as shown above. It is essential to optimize the sequencing of production in order to properly harness the available resources, avoiding as much as possible the use of overtime, hiring services and loss of orders due to lack of unavailable capacity.

### 4. Starting the production planning and scheduling with a MRP model

Enterprise Resource Planning (ERP) is one of the industrial practices employed to support financial decisions, quality control, sales forecast and manufacturing resource planning, among other fields (Gershiwin, 1986; Clark, 2003). ERP proposes an integrated solution for the whole enterprise with modules that cover areas such as: Production, Accounts, Finance, Commercial, Human Resources, Engineering and Project Management.

One ERP module is the Manufacturing Resource Planning (MRP), which supports production planning and scheduling decisions. It establishes the quantities and due dates for the items to be manufactured or assembled as well as determining the production resource needs.

The MRP model comes from a production planning perspective rather than an optimization perspective. This is a very practical model and can be used as a basis for even more sophisticated models. An interesting point in MRP is that it generates integer production quantities, provided that the demands and minimum lot sizes are integers. Computerized planning systems based on MRP have been in use for decades and its logic remains at the heart of the production planning module of many modern Enterprise Resource Planning (ERP) systems.

MRP generates the production orders to be implemented on the shop floor, working as if the shop floor had infinite capacity. It also generates production orders that sometimes overload critical resources. When overload occurs, the manager has to decide, based on his experience, which orders should be placed first, postponed and those that will not be implemented. Therefore, the MRP plan does not guarantee enough capacity to actually carry
out its implementation. This plan can also be so unrealistic that it will not be useful (Voβ and Woodruff, 2006).

The MRP works with large lot sizes. The primary reason for this is to ensure that not too much productive capacity is used to changeover from one stock keeping unit (sku) to another. But even with infinite capacity and lot size nervousness the MRP is very useful in industries with changing demand patterns where standard orders can not be used. In addition, MRP models can provide a good starting point for planning and for ordering material (Voβ and Woodruff, 2006). Although MRP is presented as a planning tool, it is also often used as a scheduling tool.

To illustrate the MRP main steps, consider product’s P structure shown in Figure 6. P has a Bill Of Materials (BOM) with three row materials (RM1, RM2, RM3), two manufactured items (M1, M2) and two assemblies (A1, A2).

![Fig. 6. Structure of a Product](image)

Each one of the components of this BOM has particular characteristics and restrictions, for instance, lead times for acquisition, lot size and initial inventory. The technological restrictions for all items belonging to the Product P BOM are summarized in Table 2.

<table>
<thead>
<tr>
<th>Item</th>
<th>Lead Time</th>
<th>Minimum Lot Size or Order Quantity</th>
<th>Components</th>
<th>Initial Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM1</td>
<td>2 days</td>
<td>Order quantity = 200</td>
<td>-</td>
<td>150</td>
</tr>
<tr>
<td>RM2</td>
<td>4 days</td>
<td>Order quantity = 450</td>
<td>-</td>
<td>450</td>
</tr>
<tr>
<td>RM3</td>
<td>2 days</td>
<td>Order quantity = 150</td>
<td>-</td>
<td>130</td>
</tr>
<tr>
<td>M1</td>
<td>1 day</td>
<td>Lot Size = 100</td>
<td>2 RM1</td>
<td>40</td>
</tr>
<tr>
<td>M2</td>
<td>2 days</td>
<td>Lot Size = 100</td>
<td>2 RM3</td>
<td>120</td>
</tr>
<tr>
<td>S</td>
<td>1 day</td>
<td>Lot Size = 80</td>
<td>3 RM2; 1 M1</td>
<td>70</td>
</tr>
<tr>
<td>P</td>
<td>1 day</td>
<td>Lot Size = 100</td>
<td>1 S; 1 M2</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 2. Technological restrictions for the components belonging to the BOM of Product P
<table>
<thead>
<tr>
<th>Component</th>
<th>Day 0</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
<th>Day 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>20</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>20</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>20</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>20</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>RM3</td>
<td>40</td>
<td>100</td>
<td>80</td>
<td>60</td>
<td>80</td>
<td>100</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>RM2</td>
<td>40</td>
<td>100</td>
<td>80</td>
<td>60</td>
<td>80</td>
<td>100</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Needs for BOM of P components
By defining the product P demand for the next 7 days as 20, 50, 40, 30, 40, 50, 30, it is possible to calculate the demands for all the items in Table 2. Knowing that the MRP is conceived to produce as late as possible, and assuming these components are used only in product P, Product P’s component needs can be determined as in Table 3.

For component P in Table 3 a), the initial inventory is 90 (as indicated by Day 0). Because of the demand of 20 units, on Day 1 the inventory plan will be 70 units (90 - 20 = 70). On Day 2 the demand will be 50 units, which will result in 20 units for the inventory plan. On Day 3 there are 20 units missing from the inventory, since the demand is 40 units and the inventory plan shows 20 units. Because of the lead time of 1 day, on Day 2 it is necessary to place an order for P1 with the amount of this order being at least 100 units, which is the minimum lot size of P.

Components RM3 and RM1 (e and g, in Table 3) have double the demand of P, because their incidence is 2 units in components M2 and M1 respectively. The incidence of Component RM2 is three to one product P. The calculations for all components are similar to those made for component P. This is a result from MRP calculations.

The amount of final products in a factory, the technological restrictions of each one of the items in their BOM, and the fact that several of these items belong to more than one BOM are issues that are considered. Thus, the MRP does an efficient job regarding the organization of dates and amounts to be manufactured, assembled or purchased of each one of the necessary items presented in all BOM.

However, this is not enough to avoid the chaos on a shop floor, since MRP does not consider the production capacity of the available resources. It assumes that each one of the resources has an infinite capacity, therefore it is always possible to start the production of an item by considering lead times only.

Besides the fact that resources have limited capacity, equipment is frequently in maintenance, or broken. These aspects must be considered by the manager at each stage of planning. It is the manager’s task to decide which order will be late, which will be outsourced and which will not be carried out.

Considering that some periods are overloaded it seems natural to move the exceed production to those periods that are less busy. But this is not a simple task, with it being necessary to use mathematical tools in order to solve this problem.

5. Mathematical model

The manufacturing process is a pre-defined technological sequence of production activities in the production network. On their way through the production system, the items, raw material, semi-manufactured products and finished products wait in queues for release conditions. They are subjected to fabrication or assembly or transportation operations until they reach the final customer. Figure 7a) represents the pre-defined sequence of production activities for one product and Figure 7b) expands it through three periods of time.

Decisions regarding the quantity of raw material available, production level for each group of machines and demand to be supplied must be made in each time period. Cost and capacity are associated with each production stage (processing, transportation, assembling and storage activities). Costs and capacities can be different for each time period, e.g. at period t, stage S, a machine M has a capacity of 3 units; while at period (t+1) the same machine M at the same stage S can have a capacity of 2 units due to a determined preventive maintenance program.
5.1 Mono-product system

Bowman (1956) was the first to suggest to solve the production scheduling problem by a transportation method framework, and further, that many transportation problem could be extended to include a multi-time period. Denoting production resources (storage and machines) by arcs and the decision points by nodes, the production routes can be represented as in Figure 7.b). RM represents the raw material storage node and the diverging arcs (RM1, RM2 and RM3) drive the raw material to each period of time. Node D represents the total demand. It results from the sum of flows through arcs D1, D2 and D3.

The flows through the arcs represent the decisions to be made. For example, figure 8 presents the balance equation for a generic node (i,t). The flow $x_{i-1,t}$, leaving node (i-1,t) and arriving at node (i,t), through arc (i-1,t) represents the amount to be processed at the production stage (i,t). Flow $y_{i,t-1}$ that leaves node (i,t-1) and goes to node (i, t) through arc (i,t-1), represents the amount of material to be stored. The decisions in the node (i, t) is the amount to be produced and the amount to stored.

For each production node the balance equation can be written as follows:

$$x_{i-1,t} + y_{i,t-1} = x_{i,t} + y_{i,t}$$

Associating cost and capacity to each arc, the optimisation problem and supplying demand at minimal production cost, can be stated as a network flow model (Carvalho et al., 1999), as shown in Problem P1.
Min $f(x)$  
\[s.t. \ Ax = b\]  
\[l < x < u\]  
(P1)  
where $l$ and $u$ are the minimum and the maximum capacity associated to a production or storage resource, respectively.

### 5.2 Multi-product system

In real production systems, several products are processed in the same facility. To handle variations, operation facilities are normally designed with sufficient flexibility to process more than one family of product. The facility capacity must be co-ordinated to guarantee that the sum of the production plan of each individual product does not exceed the production capacity of each facility.

Problem P1 therefore must be expanded to Problem P2, to represent the multi-product production system as shown in (P1).

$$\begin{align*}
\text{Min} & \sum_{i=1}^{n} c_i x_i \\
S.A.: & A_1 0 \ldots 0 0 \\
& 0 A_2 \ldots 0 0 \\
& \vdots \vdots \ddots \vdots \vdots \\
& 0 0 \ldots A_n 0 \\
& S_S S_s \ldots S_n I \\
\leq & x \leq u \\
\end{align*}$$

Where $A_i$ is the incidence matrix of product $i$, $x$ is the decision variable representing the material flow in the production resource, limited to the lower bound “$l$” and upper bound “$u$”. Demand is shown by $b$ and raw material availability vector, $S_i$ describes the mutual capacity and mutual inventory constraints, called side constraints. The production capacity vector is represented by $d$. This problem exhibits a special structure that is exploited in the solution considered by PRONET algorithm (Yamakami et al, 2000).

If the production lines are decoupled, the $S_i$ matrices are equal to zero and the optimal manufacturing scheduling is reduced to a simple network flow problem.

### 5.3 Solution algorithm

The large dimension and the very particular structure of Linear Program models associated with the production planning problem motivates the development of algorithms that explore the special structure of the problem. One approach is the Netside Algorithm (Kennington, 1980), but its efficiency is limited to the size of the problems (Carvalho et al., 1999).

Interior Point Methods (IPM) have grown in importance since the positive results obtained by Adler et al. (1989). Nowadays, the IPM for Linear Programming is well established for practical applications and good algorithms are available (Gondzio, 1996; Wright, 1996).
In addition to using the IPM for large-scale problems, faster solutions can be achieved by applying practical knowledge and special characteristics from the real problem throughout their transformation in a mathematical model. The first one is bottleneck management defended by the Theory of Constraints (TOC) (Goldratt and Cox, 1986). According to this theory, a productive system can be divided into two kinds of resources: the bottlenecks and others. Bottlenecks are those resources with limited production capacity, which therefore need special treatment. The planning and scheduling of these resources must be managed carefully in order to meet the demand requirement dates. The decisions regarding other resources are submitted to bottleneck decisions.

The second assumption is the use of mathematical transformation over the constraint matrix. According to Zahorik et al. (1984), “the immense size of these problems and the imprecise nature of many of their costs and demands further suggest that good heuristics may be as desirable as (presumably) more costly optimisation algorithms”. As a consequence of this thinking, another way to solve (P1) is to combine both techniques - optimization and heuristic - to attain the best features of each one, as suggested in the followings sections.

5.4 Integration of MRP and linear programming

MRP is an industrial practice largely employed in manufacturing production planning. It is however not enough to completely solve the complex production-scheduling problem, mainly because the basic form of an MRP considers infinite capacity. As a consequence, it does not adequately coordinate the production capacity and raw material time availability with demand requirements and requires additional tools based on expert knowledge of the problem and optimization procedure (Józefowska and Zimniak, 2008)

Temporal co-ordination can be reached by integrating a Linear Programming algorithm with the MRP module of an ERP. MPS generates long term planning for product types as shown in Figure 9.a). Starting from MPS targets, the P1 optimization program assesses the ERP Data Base considering temporal production availability, raw material constraints and product priorities, and generates feasible temporal production scheduling. The temporal co-ordination suggests, when necessary, the anticipation or postponing of demand attainment, according to a pre-defined criterion.

To overcome the dimensionality problem of Linear Programming the model only considers the important production resources and product families. It generates a temporal production scheduling for this simplified system, and due to this, its results need to be refined. MRP, starting from the targets established by the linear model, splits the families into items and explicitly considers the complete production tree and set-up time of each resource. Therefore, Linear Programming and MRP are complementary planning tools. This is because the first one has the temporal visibility of a production problem and the second one has the visibility of the items and the set-up time, and the integration of both allow for an accurate representation of the problem.

An alternative scheme is shown in Figure 9.b) that suggests running the MRP module, and temporally verifying production capacity or raw material constraint violation. When violation occurs, the optimization algorithm is executed to generate a feasible solution. In both approaches, the objective is to maximise the customer service level by co-ordinating the temporal distribution of the production according to the production availability of machines, storage and raw materials.
6. Case study

Solving real problems is not limited to the choice of which tool must be used. Each actual manufacturing system has its own characteristics that will differ from others. It is therefore regarded as unique and requires a detailed analysis of its process in order to define an adequate model. Furthermore, modeling a real situation, whether it is a factory floor, a network of product distribution or a power station, requires abstraction and approximations that transform the actual system into a well-structured mathematical problem.

At this point, the use of empirical knowledge that routinely deals with the actual situation is valuable. This knowledge can be transformed into rules that result in the computational model being as close as possible to the real situation that will be studied. The main steps transform a real problem into a computational model and can be established as:

- To understand the problem to be modelled
- To use empirical knowledge from those who works with the problem
- To transform this knowledge into well-structured rules accepted by the tool and that reflect, as much as possible, the actual situation
- To choose the tool to be used.

6.1 Stamped components industry

The above approach was applied to an Auto-parts enterprise with a commercial ERP installed. This enterprise produces wheels, chassis and stamped components for the main commercial vehicle manufacturers in Brazil, as well as exporting these products to several countries. The objective of the study was to implement a capacitated production schedule in the chassis manufacture. It has a complex production system involving several operations.

This study identified three presses (PR00, PR10 and PR20) as the bottleneck operation in chassis production. These presses with capacities of 3kt, 5kt and 3kt respectively work in parallel to one another. Each item has a predefined fabrication route (Figure 10) with one or two operations in the presses, called operations RF and EF. The 5kt Press can process any item, but items I1, I3 and I4 can be processed only in this particular Press. Since the processing time of each item is around 0.01 hour, while the set-up time for an operation varies from 0.01h to 2.0h, it is interesting to process the largest possible number of items when set-up time is performed. Therefore, a request for an item in a week should be processed as a complete lot.

Practical operations have shown that the processing sequence of items can influence the system set-up time. A heuristic method that aggregates items into families was implemented.
to take advantage of this particularity. The heuristic method is based on the fact that the setup time of the press operation includes the removal, arrangement or addition of blocks required to process each item in the press. One of these operations will result in a new product.

![Fabrication Routes in the Presses](image1)

**Fig. 10. Fabrication Routes in the Presses**

Figure 11 shows the steps for the process preparation sequence within a family, considering that the processing sequence is determined by the number of blocks needed to process each item in the family. The family item with the largest number of blocks starts the process. A new set-up is then performed and the item with the second largest number of blocks of this family is processed.

The heuristic rules to schedule items I141, I142 and I151 are: start with the item in the family that needs the largest number of blocks. After this item is processed, the next in the sequence is the one that needs the second largest number of blocks.

The heuristic method for each family is:

- **Step 1.** Identify the item with the largest number of blocks;
- **Step 2.** process the lot of this item;
- **Step 3.** remove, arrange or replace the necessary blocks to process the next item with the largest number of blocks. Repeat steps 1 to 3 until the last item of the family has been processed;
- **Step 4.** select a new family and go to step 1.

![Items processing sequence](image2)

**Fig. 11. Items processing sequence**
When the processing sequence follows the heuristic method, the set-up time for items of the same family varies from 1 minute to 1 hour and 30 minutes, while set-up time between families is 2 hours. Figure 12 shows a comparison between set-up times for the same group of items. In the first case, items are processed following a FIFO (first in first out) sequence. In the second case the heuristic method grouped the items into families.

![Fig. 12. Set-up times with and without heuristic method](image)

Figure 13 shows the relation between set up time and process time for aggregation of items into families. In this figure items 1 to 4 represent the first family. The second family is made up of items 5 to 8. The set-up time for the first item processed into a family is 2 hours. It is therefore possible to observe that 1, 5, 9, 11, 13, 15, 16, 17, 18, 19, 21 and 23 are the first items in their families, requiring 2 hours for set-up. Families 15, 16, 17 and 18 are composed of just one item.

![Fig. 13. Families processing sequence](image)

Tables 4, 5 and 6 present actual auto-parts industry data. The first column in Table 4 identifies the families; the second column: items. The third column presents the performed operation – called RF or EF; the fourth column presents the best process sequence for items.
within a family and column five specifies possible equipment to be used. The suggested sequence in column four was obtained by following the heuristic method.

This table presents two families and their operations in the presses. Four items from family F1 are submitted to RF and EF process operations are (I161, I171, I121 and I131). Three are submitted only to EF process operations (I141, I181 and I191). Four items from family F2 are submitted to RF and EF process operation (I241, I281, I272 and I242) and two are submitted only to EF process operations (I282 and I251).

In daily operations, demand for the complete family of products is unlikely to occur. However, even when there is demand for part of the family the pre-established processing sequence must be observed to minimise the processing time. For example, suppose that in a week the factory had a demand of I272, I241, I171, I121 and I242. The ideal processing sequence should be as shown in Table 4, which means, for Family F1, the item I171 is followed by item I121. For family F2, the sequence should be: I241, I242 and I272 for RF operation and I241, I272 and I242 for EF operation.

<table>
<thead>
<tr>
<th>Item</th>
<th>Operation</th>
<th>Sequence</th>
<th>Press</th>
</tr>
</thead>
<tbody>
<tr>
<td>I161</td>
<td>RF</td>
<td>1</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I171</td>
<td>RF</td>
<td>2</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I121</td>
<td>RF</td>
<td>3</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I131</td>
<td>RF</td>
<td>4</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I161</td>
<td>EF</td>
<td>1</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I171</td>
<td>EF</td>
<td>2</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I131</td>
<td>EF</td>
<td>3</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I141</td>
<td>EF</td>
<td>4</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I121</td>
<td>EF</td>
<td>5</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I181</td>
<td>EF</td>
<td>6</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I191</td>
<td>EF</td>
<td>7</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I241</td>
<td>RF</td>
<td>1</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I281</td>
<td>RF</td>
<td>2</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I282</td>
<td>RF</td>
<td>3</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I242</td>
<td>RF</td>
<td>4</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I272</td>
<td>RF</td>
<td>5</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I251</td>
<td>RF</td>
<td>6</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I241</td>
<td>EF</td>
<td>1</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I281</td>
<td>EF</td>
<td>2</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I272</td>
<td>EF</td>
<td>3</td>
<td>PR00/PR20</td>
</tr>
<tr>
<td>I242</td>
<td>EF</td>
<td>4</td>
<td>PR00/PR20</td>
</tr>
</tbody>
</table>

Table 4. Best process sequence inside a family

The first column of Table 5 is dedicated to families, the second to operation. The fifth column presents the set-up time for exchange items from column three to items in the fourth column. Suppose item I242 is leaving the press in the first operation. Considering the ideal sequence presented in Table 4 for the RF operation, the next item to be processed is I272 with a set-up time equal to 10 minutes - Table 5. For the EF operation, the ideal sequence will be I272 followed by I242 and the set-up time will be 1 hour and 30 minutes.
Table 5. Best process sequence within a family

Table 6 exemplifies the demand for items in 2-week periods.

Table 6. Demand per period

Two algorithms have to be developed to transform the data from ERP format to optimizer format: “Families” and “Network”, Figure 14. The “Families” module works on the data provided by Table 4 to Table 6 and on the ERP data base (like sequences and bill of material) to create “equivalent families” - EQ, which are the subsets of original families.

The demand size is considered equal to a unit for each EQ. Its processing time is the sum of the processing time of each item with demand in the period plus set-up time (ERP data base and tables 2 and 3).
Table 7 illustrates the procedure for the creation of EQ. In the first column, the families are presented, the second column contains the operations to be processed (RF or EF), and the third column presents the items with demand during the period. The fourth column contains set-up times for each item in the family. The first item of each family always needs 2 hours set-up time, in the first or second operations. The set-up time between the first and second item in family F1 at RF operation is 30 minutes and at the EF operation is 40 minutes. The fifth column shows the sum of set-up times of the RF and EF operations for each family. Column 6 shows the demand during the period. The processing time is shown in column seven (50 minutes is necessary to process 70 items of I171). Column eight contains the sum of processing time for the whole family in operations RF and EF. Column nine shows the Equivalent Family name, and column ten the time needed to process the family. E.g.: F1A processing time is (5 hours and 10 minutes + 3 hours and 30 minutes = 8 hours and 40 minutes).

<table>
<thead>
<tr>
<th>Fam</th>
<th>Op.</th>
<th>Item</th>
<th>Set-up (h:min)</th>
<th>Total Set-up (h:min)</th>
<th>Dem (item)</th>
<th>Proc. Time (h:min)</th>
<th>Total Proc. Time (h:min)</th>
<th>EQ</th>
<th>Proc.+ Set-up (h:min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>RF</td>
<td>I171</td>
<td>2:00</td>
<td>5:10</td>
<td>70</td>
<td>0:50</td>
<td>3:30</td>
<td>F1A</td>
<td>5:10+3:30 =8:40</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>I121</td>
<td>0:30</td>
<td></td>
<td>85</td>
<td>1:00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EF</td>
<td>I171</td>
<td>2:00</td>
<td></td>
<td>70</td>
<td>0:45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EF</td>
<td>I121</td>
<td>0:40</td>
<td></td>
<td>85</td>
<td>0:55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>RF</td>
<td>I241</td>
<td>2:00</td>
<td>7:50</td>
<td>50</td>
<td>0:30</td>
<td>3:50</td>
<td>F2A</td>
<td>7:50+3:50 =11:40</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>I242</td>
<td>1:30</td>
<td></td>
<td>80</td>
<td>0:40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>I272</td>
<td>0:10</td>
<td></td>
<td>80</td>
<td>0:45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EF</td>
<td>I241</td>
<td>2:00</td>
<td></td>
<td>50</td>
<td>0:30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EF</td>
<td>I272</td>
<td>0:40</td>
<td></td>
<td>80</td>
<td>0:45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EF</td>
<td>I242</td>
<td>1:30</td>
<td></td>
<td>80</td>
<td>0:40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Equivalent family creation procedure

Since EQ is seen by PRONET as one item, this guarantees that all items in the family are processed together, thus respecting the best heuristic sequence. This approach enables the use of Linear instead of Integer Programming, overcoming the dimensionality and time consumption problems.

6.1.1 Example

The past production data of the company identifies around 900 items that are grouped into 92 families, organised as in Tables 4 and 5.

Table 8 considers an extract scenario of two periods. Each family and processing time are similar to those presented in table 4. Column 1 shows the Equivalent Family name. Column 2 shows the original MPS schedule for 2 periods. The first 6 lines are related to the first period, and the 11 remaining to the second period. “Press” columns refer to the equipment used to process the family. In this way, family F20A must be processed in press PR20 in the first and second periods. The choice of equipment respects the manager’s priorities.
Table 8. Percentage of families and respective amount of time processed by period

Columns 3 and 5 indicate the percentage of demand of each family to be processed in each period. Columns 4 and 6 give the amount of processing time in each period. Therefore, family F6A, initially allocated to MPS for processing in the second period, will have 32% of its production, or 18.5 hours, advanced to the first period. F20A and F12A will also have part of their production advanced to the first period.

If the press has not enough capacity to process a family, a heuristic rule that inverts equipment priority, is used. F6A is programmed for the second period, in press PR00. This press does not have enough capacity to process the whole family. Another possible press in this case is PR20, which, during the second period, has the capacity to process only 68% of the family.

Software suggests advancing 32% of the family to the first period, in the prior press – PR00. Table 9 presents the total amount of time spent in each press in both periods. Analysing this table, the manager can decide, for instance, to process F6A in press PR20 in the first period.

Table 9. Work load in presses in two periods

The example presented in this work is the extract of a simulation for 40 families in 4 periods.

In this example some products can be classified as one-stage production scheduling problem in which each activity requires two operations to be processed in a bottleneck stage (resource). A more complex problem is that one in which a product each activity requires two operations to be processed in stages 1 and 2, respectively. In this case there are two
options for processing: the first is to produce it by utilizing in-house resources, while the second is to outsource it to a subcontractor. For in-house operations, a schedule is constructed and its performance is measured by the makespan, that is, the latest completion time of operations processed in-house. Operations by subcontractors are instantaneous but require outsourcing cost. Computational model for and approximate algorithm for this NP-hard case can be found in Lee and Choi (2011).

6.2 Auto part gears industry

a. Enterprise characteristics and objectives

A medium-sized manufacturing enterprise with around 200 employers produces auto part gears through 17 cell manufactures numbered from 1 to 25. In general, each cell is dedicated to manufacturing a fixed set of products expressed by a correspondence table, as shown in Figure 15. However, in daily operation there may be a product being processed in another cell, due to a lack of capacity of the priority cell. However, this operation causes disruption in production and should be avoided whenever possible.

The production of gears is determined by the client’s requests. The company works with both regular and special items. For some regular items, the company has a forecast of more than one year, discretized weekly. For others there is forecast of eight months. In the case of special items, the demand is small and irregular. The cells work in two shifts of 8.5 hours per day with a monthly limit of 360 hours.

Currently the setup time is the time between item entry in the cell and its exit from the last operation. As in the gears process, the setup time of the cells are time consuming (on average 4-5 hours). They are performed concurrently with processing the previous batch, i.e., as the last part of a batch releases one machine of the cell, this machine is immediately prepared for the next batch. In general, this operation is initiated by an operator and completed by the preparer.

![Fig. 15. Products X Cell correspondence](image)

The company has a homemade MRP supported by Excel spreadsheets® to control shop floor operations. This MRP, as well as in any of the MRP’s encountered on the shelf, does not consider the limits of resource capacity and its method of production programming does not support the decision maker with a temporal production view. In order to support the decision maker in the difficult task of optimally allocating the production lots through the 17 cells along a specified time period, given the quantity demanded and the dates requested, an optimization model based on graph theory was developed. The objective of this graph is to minimize storage and backorders. Taking into account both the availability in the company of the spreadsheet, as well the internal knowledge of this tool, a standard environment for both modelling and presentation of results was developed.
b. Spreadsheet model

As it is a medium-sized enterprise, an excel worksheet was suggested as an environment for the solution, as pointed out in the table 10 whose main function is:

Table 10. Auto part gears spreadsheet model.
The objective of this implementation is to show how simple it is to develop a model and to develop a schedule using a spreadsheet. The results table shows the scheduling for each product and cell at each time period. It also presents the backorders of 470 units and 1000 units for the products P48010 and P48011 respectively in the period 2.

7. Conclusion

Production scheduling involves operation propositions for real complex systems and despite the ultimate objective of producing a schedule that satisfies domain constraints and optimizes overall performance, scheduling, in the most practical domains, is concerned with solving a problem of a much larger scope involving specification, negotiation and refinement of input requirements and system capabilities. Production scheduling implies the development of a set of activities that lead to most effective overall system performance through simultaneously balancing resource availability, enterprise interests, suppliers constrains and contracts and client needs.

It is therefore unreasonable to expect to fully automate the analysis process in a single system that covers all real applications. The search space is unwieldy and ill-structured, and human expertise is needed to effectively direct the search process. At the same time, problem scale generally demands substantial automation. The research challenge is to offer the user a flexible environment and to incorporate it into the scheduling process, without requiring understanding of the system’s internal model. In other words, the system must bear the burden of translating to and from user interpretable representations. It is important to convey the results in a form that facilitates comprehension and conveys critical tradeoffs, and to accept user guidance on how the system model can be manipulated.

Taking into account these aspects, this work used three different approaches to solve real problems: Mathematical Programming (Linear Program), Industrial Practice (ERP) and Manager Expertise, translated into heuristic rules. Mathematical Program results will give a “near optimum” value, since calculations are made based on estimated set-up times aggregated to the processing times. Better representation of set-up times is reached through the MRP module. Knowledge of shop floor particularities allows the manager to identify characteristics that improve the performance of the production system and that can enrich the solution algorithm. The production optimisation balances the capacity with temporal demand oscillation and induces, when necessary and possible, transferring production from overloaded to idle periods. This in turn avoids extra hours and contracting outsider services and to reach the desired production level by advancing or postponing production to idle periods.

The advantage of this integration is clear. It helps the manager deal with multiple approaches added to enterprise competences to support his decisions overcoming the usual disruptions by radical changes in enterprise procedures.
Although this chapter focused on the creation of mathematical models to be coupled with traditional ERP packages there are however numerous alternative model forms and Pidd’s work provides considerations of a wider view of modelling (Pidd, 2003).

New models and methods have to be coded for each new real problem. One way to achieve this is the use of modelling language that provides algebraic and set notations to concisely express mathematical problems that can then be solved using state-of-the-art solvers. These modelling languages, namely AMPL (Algebraic Modelling Language for Mathematical Programming), do not require specific programming skills and provide flexibility to represent production planning models to any enterprise.

8. References


Gondzio, J. 1996 Multiple centrality corrections in a primal dual method for linear programming, Computational Optimization and application, 6, 137-156, ISSN: 0926-6003.


Generally speaking, scheduling is the procedure of mapping a set of tasks or jobs (studied objects) to a set of target resources efficiently. More specifically, as a part of a larger planning and scheduling process, production scheduling is essential for the proper functioning of a manufacturing enterprise. This book presents ten chapters divided into five sections. Section 1 discusses rescheduling strategies, policies, and methods for production scheduling. Section 2 presents two chapters about flow shop scheduling. Section 3 describes heuristic and metaheuristic methods for treating the scheduling problem in an efficient manner. In addition, two test cases are presented in Section 4. The first uses simulation, while the second shows a real implementation of a production scheduling system. Finally, Section 5 presents some modeling strategies for building production scheduling systems. This book will be of interest to those working in the decision-making branches of production, in various operational research areas, as well as computational methods design. People from a diverse background ranging from academia and research to those working in industry, can take advantage of this volume.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: