Chapter 1

Introduction

1.1 The scheduling problem

The scheduling of batch processes is one of the most complex and important problems faced by a wide variety of processing industries. Some early works (Lazaro, 1986) estimate the production of the batch processing industries as 60% of the total products generated in the chemical industry worldwide and represents the 80% of the products associated to the pharmaceutical industry (Reeve, 1992). In spite of this importance, recent papers (Honkomp et al., 2000; McCall, 2001) remark that presently scheduling is often a manual procedure, which leads to an operation characterized by high inventories, inefficient capital utilization and increased operation costs. There are also reported complains about the lack of powerful, easy-to-use, PC based tools, which can solve detailed operational problems, as well as perform high level analysis across the supply chain.

This problem has been the focus of an important number of papers in the recent years, especially from the early 1980's to nowadays although the industry has been interested in effective ways of solving the scheduling problem since the early 1940's (Shobrys, 2001). An extensive work has been done as it is shown in reviews like in Shah (1998), but the complex nature of the scheduling problem, defined as a NP-complete problem, results on the lack of an unique solution widely used in the industry as shown by Pekny and Reklaitis (1998).

The planning and scheduling problem is basically a decision-making problem, which ends up to a set of decisions as:

- How much to produce of each final product and intermediates.
- The unit and storage assignments, the definition of what units are producing each of the intermediates and final products and where to store them.
- The timing of all the operations involved in the production of all the products.
- The use of other required resources available, as manpower, steam, etc. necessary to execute all the operations scheduled.

A good decision-making in the previous aspects provides an improvement in the efficiency of the plant operation in the following ways:

- Increased plant capacity avoiding dead times. A good production schedule eliminates bottlenecking situations in advance, so increases the plant capacity allowing more of each product to be produced in a given planning period.
- Better use of equipment, manpower and other resources. This is the result of the unit and resource assignment. The better utilization of equipment and manpower in a particular problem can even help to reduce unexpected downtimes by allowing for planned maintenance at approximate times in the production schedule.
- Cost/benefit trade-offs for new scenarios. A good scheduling allows to perform an analysis of the influence on the whole process of the decisions taken and how the inclusion of new products and new investment in production resources will affect to the general performance of the plant.
- Improvement in the use of intermediate storage. The scheduling decisions also affect the use of intermediate storage. A good storage management is needed to keep the intermediate products inventory to a minimum, thus avoiding high inventories and the additional cost related to the storage of an excessive amount of intermediates.

The benefits derived for the implementation of an scheduling system in the industry are clear. For instance, one case study (Thomas, 1999) related with the implementation of an scheduling system in a pharmaceutical industry shows the following impact issues:

- The planning procedure required fewer personnel (from 120 men/hours to 40 men/hours per week)
- The planning horizon was increased from one week to three months
- Product lead time was reduced from 16 weeks to 2 days
- Forecast accuracy at the product family level rose from 55% to 96%
- Forecast accuracy at the major product level rose from 0% to 99%
- The project paid for itself within three moths of go-live.
- Capacity exceeded demand

Other sources (McGettrick et al., 2001) show other benefit figures applied to a broader range of industries but the overall conclusion is that the project pay-back for the implementation of a planning and scheduling system can often come within the first year of application.

1.2 Scheduling requirements for industrial problems

In the paper of Honkomp et al. (2000) there is an extensive review of the common industrial requirements for the scheduling problem. In the industrial experience is common to find dozens of specific details and plant policies that are usually difficult to be included into mathematical models. The reality is much more detailed and complex than the mathematical models may be, so the last are forced to include strong simplifications. The excessive simplification often results on optimum schedules which are infeasible in practice, so a fair amount of changes in the schedule by the planner may still be needed to send a schedule to the shopfloor due to the limitations of the model used.

Hence, it is important to have a good description of the scheduling requirements usually found in industrial problems in order to identify the real needs of the industry. Basically, there are six main points involved in any industrial problem:

- Robust problem definition
- Problem Size
- Shared Resources
- Storage of materials and products
- Non-productive activities
- Demand management

This classification of needs according to Honkomp et al. (2000) provides an industrial point of view of the scheduling problem. This industrial point of view is essential if the objective is to provide industrial-useful solutions to the scheduling problem.

1.2.1 Robust problem definition

There are three critical aspects in the robustness of a solution of a scheduling problem of the process industry: the objective function, the uncertainty in the future production scenario and the time description.

The maximum production at the minimum cost is usually the main objective for the industry. Nevertheless, it is difficult to measure and identify all the costs related to the process. Moreover, the operation in a plant is of dynamic nature, an optimal solution valid in one moment can be clearly infeasible some time later. This situation can be avoided providing a more robust schedule. The objective function should be flexible enough. The basic objective is to find a feasible answer, but could be much better if a solution close to the optimum, feasible and robust is obtained. So a good definition of the objective function that takes into account not only the profit but also the quality of the schedule obtained is needed.

Industrial processes are often running with operating times that can vary under unpredictable parameters (i.e. the weather conditions). This aspect generates a difference between the predicted and the realized schedule. This offset is usually corrected performing periodical rescheduling. If the predicted schedule is not robust enough, the rescheduling should be performed more often. The key aspect of a robust schedule is to obtain a schedule with degrees of freedom that absorb the possible variation of the production times without sacrificing the value of the objective function.

The industrial production works on a continuous time basis. A continuous time description in the model used is required to find feasible schedules in the industry. Coarse definitions of time are not useful if the details in the process are important as usually are.

1.2.2 Problem size

The industrial problems tend to be large when considering all the aspects involved. Several aspects contribute to increase the size of the problem such as the number of materials involved, the number of processing units, the time horizon and the time dependent inputs:

- The number of materials (raw materials, intermediates and final products) can easily be in the order of hundreds and sometimes up to several thousands.
- One product can usually be produced in different processing units, even with different processes and can be packed in several different ways. This aspect increases the combinatorial complexity of the problem.
- The number of processing units can also be large, specially if it is necessary to track all the auxiliary equipment as pumps or if multiple sites are scheduled simultaneously.
- The inputs can vary over time, new commands and new products can be added to the process increasing the size of the problem. In fact, the industrial problem is not static, several conditions can modify the problem along the time.
- The time horizon is usually in the order of weeks for process scheduling, months for planning and years for design. As bigger is the time horizon, bigger is the problem to be solved if the full detail of the solution is maintained.

1.2.3 Shared resources

The use of shared resources between processes increases the complexity of the problem. The constraints associated to the shared resources can delay the start time of one operation. In other terms, limits in resource shared between two operations can increase the processing time of one of them or both. This last point is complex as the length of one operation is resource-dependent and cannot be known exactly a priori.

The use of shared resources is an important and common aspect in the industrial problems. The typical shared resource is a utility as steam or electricity, but there are many more situations in the industry that fit in the definition of shared resource.

One example of shared resource is equipment that has to be shared between different operations. This equipment can be, for example, cleaning equipment necessary to perform the cleaning tasks. The use of this equipment forces an interaction between different batches. The problem is harder to solve if the shared equipment requires to be used for a long time in the different batches.

Another typical shared resource is manpower. In some operations the presence of operators is required. Additionally, each operation requires an operator with certain skills. The introduction of the operators in the problem increases the complexity, as the availability of the operators depends on several factors as: the number of operators present in each shift, the possibility of overtime, the possibility of flexible crewing under certain limits and the skills offered for each operator in each shift. It is clear that some operations can not be performed without the presence of an operator, forcing the task to be delayed.

Sometimes, several production lines use common equipment. For example, units which are fed from a shared bulk tank. In this case the units which share the bulk tank can only produce similar formulations. This situation affects the scheduling as all the units that share the resource are forced to produce a similar product.

Other utilities such as cooling and heating can affect to the production in several ways. There are two main situations:

- An specific amount of cooling or heating is required and the operation cannot be started unless an specific amount of the heating (steam) or cooling is available.

- The operation can get started without all the required amount of the available resource. In this case the operation time will be increased, for example cooling or heating times extended. This aspect is very difficult to be included in a mathematical model as the operation time cannot be estimated a priori.

Another case of variation in the operation time can occur related to the use of labor resources. There are some tasks, i.e. maintenance task or changeovers, which duration depend directly on the number and skills of the operators available to perform such task. This situation can be observed typically on processing industries that do not have the ideal number of skilled people to perform certain tasks.

1.2.4 Storage of materials and products

The main function of intermediate storage is to decouple different operations. However, the intermediate storage possibilities increase the complexity of the scheduling problem.

The product stability can be a key issue for the use of storage. This aspect forces a limit on how long a product can be stored either to be used in another process or to be delivered.

Sometimes a minimum storage time should also be considered to allow enough time to perform the quality insurance tests. So the storage time can be limited with a maximum and minimum storage time.

The number of intermediate storage tanks used may be limited. A good assignment of the use of these tanks should be performed to avoid situations where an intermediate is produced before it is needed, blocking the use of a tank. When the number of tanks is limited causing a bottleneck, these kind of situations should be avoided.

Tanks can be also used in the production recipe as mixing units. This means that storage tanks can have the functionality of an unit that can perform different operations in a recipe.

Layout and use constraints should also be considered, specially if the charge and discharge times are long (i.e. big quantities transferred or low flowrate because viscosity, etc.) In this case special constraints should be applied depending on the material transfer capacities of the storage tanks and associated elements.

1.2.5 Non-productive activities

These kind of activities are as common as important in all the processing industries. There are several topics to be considered like setup times, good manufacturing practices, maintenance tasks, cleanings and even related to human factors, like the lunches and meetings.

The equipment setup can be dependent of the sequence of products to be processed in a specific unit. On the other hand, the setup time can only be dependent of the next product to be processed. In any way, setup times have to be considered.

Some Good Manufacturing Practices (GMP) are also examples that fit into the non-productive activity. In some industries, specially the food and pharmaceutical sectors the equipment cannot run for more than a certain number of hours without an activity to eliminate microbial material. This clean-up should be taken into account in any schedule.

Cleaning tasks should also be performed when an equipment has been idle for longer than a defined hours in order to prevent biological contamination in the products.

Another key issues are the maintenance tasks, which are usually a function of their running time. As much time the equipment is used as much frequent are the maintenance tasks. The maintenance tasks can also be scheduled to best fit the general objectives of the plant.

Other aspects related with the non-productive activities (i.e. limited capacity of the wastewater treatment plant) should also be considered as the changeovers, product purges and cleaning have effects on the amount of wastewater typically produced in these operations.

1.2.6 Demand management

The demand management is usually coupled with the requirements for raw materials, intermediates and finished products. In all the industry it is necessary to know how much and when the raw materials will be required for production.

Several demands can be satisfied producing an unique batch of product. This is known as batch splitting. This situation is common if the manufacturing of a certain intermediate is necessary to produce the final products. Several final products may require the same intermediate, but all the intermediate required can be manufactured in a single batch.

In the same way, a product cannot be manufactured before the availability of the intermediates required for its production. Therefore the quantification of the amount of intermediates required and the scheduling of the production of the intermediates should also be taken into account.

The scheduling of the shipment of the final products as well as the arrival of the raw materials should also be considered, as the capacity for those shipments is also limited and tightly related with the production schedule.

Sometimes it is infeasible to fit all the due-dates of a list of demands. In this cases, a certain degree of flexibility and secondary objectives that can be very variable are needed. There should be a decision methodology to determine what shipment should be out of date, this can be done usually using a priority or penalty system which determines what demand is less important and could accept a certain degree of flexibility.

The campaign concept can also be important, specially with products that have a seasonal demand and a long production time (i.e. beer: the orders are basically daily but the production should be started more than a month before). In this case, the schedule is not so tied to the demand, but to the campaigns or seasons. A good forecast of the orders in a mid-term horizon is important as an input to the planning and scheduling tool.

Other information like specific product sequences can be also important in the optimization of a schedule. There are certain sequences of production that can lower the production cost due to a less cost in the changeovers between products. This information should also be included in the scheduling problem, specially if the cost of the changeovers can be high.

1.3 Objectives of this thesis

As it has been shown, the scheduling problem on the industry is highly complex. The main objective of this thesis is to provide a general framework that, at this point, is developed to deal with some of the specific problems of the industry but that can be easily extended in future works to cover most of the industrial requirements. Some of the most relevant aspects to be considered are:

- Development of a new framework which allows the description of complex temporal relationships within the recipes and the different batches in a schedule.
- Calculation of the use of different resources implied and application of resource constraints.
- Application of calendar constraints in the use of resources and units.
- Detailed modeling of the storage constraints.
- Assessment of different sequencing and assignment methodologies.
- Assessment of general optimization methods.

1.3.1 Outline of the thesis

The introductory chapter is followed by a review of the relevant literature in the subject of this Thesis (second chapter). The different approaches available are presented and discussed.

A general data model for the scheduling problem in the industry is presented in chapter three. This data model is generic and can be easily extended to include additional information in future work. The detailed timing model developed in this thesis is presented on chapter four. Resources constraints and storage constraints are also presented. Two different ways of solving the timing problem are presented and compared. On chapter five, several solving strategies are presented. Sequencing, assignment and rescheduling methodologies are described. Chapter sixth is dedicated to optimization. Stochastic and exact methodologies are applied in the context of the general framework presented in the previous chapters. On chapter seven, several industrial case studies are presented and solved using the methodologies and tools described in the previous chapters.

The tool that integrates all the methodologies described in this Thesis is presented in annex A.

Finally conclusions and possible future lines of research are detailed in the last chapter.

Chapter 2

State of the art

2.1 Introduction

Planning and scheduling techniques in the process industries have been developed since the early 1940’s (Shobrys, 2001) even before the first computer appeared. In recent years, specially from the early 1980’s, this problem has been the focus of a substantial number of works as stated in the extense review of Shah (1998). The importance of the scheduling in the industry was shown in the mid 1980’s when major chemical companies realized that they were becoming limited in their ability to offset decreasing profit margins with improvements to the manufacturing process, and started examining their supply chain activities. BASF, DOW and Du Pont started to use in house developed products for their planning and scheduling. The economical advantages of the implementation of and scheduling system into a company soon become clear (Hess, 2000).

The problem of what, where and how to produce is recognized as a complex problem that implies different layers of detail through the whole production process. The different layers involved in the production process are described in detail in both S88 and S95 ISA standards (International Soc. for Measurement and Control, 1995, 2000) and shown in figure 2.1.

The highest level is used to perform the coordination between different sites. The sites can be either of the same company or even of different companies (i.e. suppliers and clients). This kind of problem is very common in the automotive industry where there are several standard ways to communicate the requirements and due-dates between the different companies involved. This level of information is out of the scope of this thesis.

The single site production planning and (on-line) scheduling are the following levels of detail in the process operations hierarchy. This work in focused in these levels of information. The site production planning is concerned in how to meet the production targets in the production site (at least try to get a good enough solution if some targets cannot be meet). Basically, it involves the details of resource allocation for each batch to be produced as well as how many batches of each product should be produced. The important information here is the sequence in which resources should be used along the production.

The on-line scheduling level is concerned in how to predict the starting and finishing times

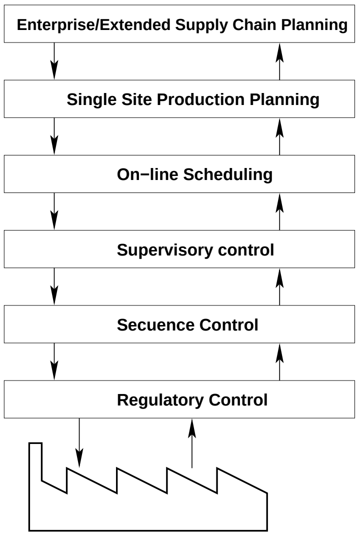

Figure 2.1: Process operations hierarchy

of the required operations according to the actual state of the plant. This level of detail is able to assure that the predicted schedule is feasible as it involves a greater level of detail as well as it uses the real information coming from the plant. At this level can be also compared if the real schedule deviates from the predicted one and how to perform changes to the predicted schedule (rescheduling).

The lowest levels of the hierarchy are related with the processing of the schedule into phases and the dispatching of the phases to the lowest level of plant control. The supervisory control is also the responsible of generating the historical of the already produced batches and informs to the upper levels of the current state of the plant. These levels of detail are out of the scope of this thesis.

The integration and the study of the information flows between the different levels of the process operations hierarchy is a common topic in the recent research (Nowicki and Brandl, 2000; Shobrys and White, 2000) and also a common point of interest of the industry, specially in the area of supply chain (Kafoglis, 1999; Nowicki and Brandl, 2000; Ferrar, 2001a; McGettrick et al., 2001) .

2.2 Single-site Scheduling

The work done in this thesis is addressed to the single-site scheduling problem down to the detail required to perform on-line scheduling in an effective way. As noted in the previous chapter, the scheduling problem has a high degree of complexity. It is usually concerned with specific production requirements (customer orders, stock requirements, long-term planning..) and with resource assignments of each operation needed to meet those requirements.

Operations can be either process operations (i.e. reaction, blending, packaging, distilling, etc.) as well as other activities (cleaning, changeovers, maintenance, quality control, etc.) There are also external time requirements as due-dates (the manufacturing should meet the expected customer orders).

The assignment of the available resources to the different operations must be carried out in the most efficient way. The most efficient way is defined by means of an objective function. Typical objective functions are the minimization of the makespan, minimization of tardiness, maximization of profit, etc.

One of the reasons of the complexity of the scheduling problem is that it is NP-complete (Pekny and Reklaitis, 1998). This means that, in the worst case, execution time for an exact algorithm on a given scheduling problem is exponential in some measure of the problem size. Therefore unacceptable long calculation times could be necessary to perform the exact calculation of the optimal solution of a scheduling problem. According Pekny and Reklaitis (1998) there are three approaches to deal with the scheduling problem:

1. *Change the problem to be solved to make it easier*: This is the most popular strategy for addressing the NP-complete nature of scheduling problems. It usually involves simplification in the structure of the recipes coarse definition of the time (discretization), ignoring important constraints such as labour, etc.

2. *Use an exact Algorithm*. This approach basically consists in trying to solve the scheduling problem using mathematical programming. This is by far the most common approach in the recent research works regarding scheduling. The NP-completeness characteristic of the problem causes that an exact algorithm may require unreasonable time to solve an scheduling problem. As Stated by Sinclair and Zentner (1999) "*Because of NP-Completeness, it is virtually impossible to construct, using MILP or any other method, a general-purpose combinatorial problem-solver with an algorithm that will work in all situations*". To deal with this long solution times usually simplifications are taken and also customized search algorithms are used in order to obtain solutions in reasonable computational times. One main problem of this kind of approach according to Honkomp et al. (2000) is that "*assumptions are made which ultimately yield solutions that are operationally infeasible*".

3. *Use a heuristic/stochastic algorithm*. This approach is widely used in available commercial scheduling tools. This kind of approach usually uses assumptions to make easier the problem to be solved. The main advantage is that they have always reasonable computation times, but on the other hand, they do not assure to find the optimal solution nor to find a "feasible"[1] solution according to Pekny and Reklaitis (1998).

2.2.1 Heuristic/stochastic approach

The heuristic approaches are concerned with formulating rules for determining sequences of operations. It is usually considered that they are best suited to processes where the production

[1]Understanding feasible solution as the one which meets all the imposed resource limitations, manufacturing constraints and product requirements. Usually this kind of methods may offer solutions that do not meet all the requirements i.e. the due-dates. In contrast the exact algorithm would fail to obtain a solution if the due-dates can not be met by a "feasible" solution.

of a product involves a preespecified sequence of operations with fixed batch sizes. It is often assumed that fixing the front-end product sequence will fix the sequence of activities in the plant. This kind of assumption is called the “permutation schedule” assumption. Normally the heuristics involve two steps:

1. Pre-ordering work of the batches by due-date, processing times , etc.

2. Place the different operations involved in the order or priority given by the previous step.

These kind of heuristics are basically coming from the job-shop scheduling problems as stated by Lazaro (1986) and Kuriyan and Rekalitis (1989). Basically these kind of algorithms have two different parts, the sequencing rules and the assignment rules as showed by Lee and Kim (1999). The combination of the different rules has also been studied for the job-shop scheduling problem as shown by Yoshida and Touzaki (1999). Other authors have developed specific rules for the flow-shop problem, as the LSL (Lowest Storage Level) distpaching rule developed by Graells et al. (1996). Other algorithms have been successfully applied to case studies that involve a large number of operations (Lee et al., 1998) and compared with the results obtained using mathematical programming models.

Most of the heuristic methods are focused on the discrete manufacturing industries and some of them, by its own definition, are not applicable to the flow-shop industry (i.e. OTOPT rule, which selects the operation with the smallest ratio of operation processing time to processing time of a single part) as there is no production of discrete parts involved. Other problem of using the classical distpaching rules is the batch size. In discrete manufacturing the batch size can vary in a quite big range whilst in the flow-shop industry batch size is usually associated to the size of the equipment assigned to the batch. This is specially important as a change of the processing unit can automatically force a change in the batch size.

An improvement over the heuristic rules is the use of stochastic search algorithms. These algorithms are evolutionary procedures that improve a given solution by modifying it looking for a better solution at each step.

One of the first stochastic search method applied to the batch scheduling problem was the Simulated Annealing (SA), studied by Ku and Karimi (1991) who conclude that SA has 90% probability of finding an optimal solution applied to the multiproduct case. In a later work by Graells et al. (1996), SA was also applied to the multipurpose case under storage constraints. A more recent work (Park and Jung, 1999) also applied a modified SA procedure in order to find an optimal schedule. One of the advantages of the SA is that is an algorithm very simple to implement and can improve solutions obtained by heuristics. However, due to the nature of the SA, it is impossible to guarantee the optimum solution in a short calculation time. A more detailed explanation of SA and how it is applied to the work done in this thesis is explained in chapter 6.2.1.

Other optimization technique used in scheduling problems is Genetic Algorithm (GA). Genetic algorithms have been successfully applied to the scheduling problem by Jung et al. (1998) where the results of applying GA are compared with the results of SA and an heuristic algorithm, the RAES (Rapid Access Extensive Search). In that paper the GA performs the best results finding the optimal solution in most of the case studies showed.

In the work of Azzaro-Pantel et al. (1998) GA are also used, in this case also using a discrete event simulator to evaluate the fitness of the different members of the population. They concluded that the use of GA represents a good alternative for treating large combinatorial problems.

Other recent works, as the one of Wang et al. (2000), also use genetic algorithms to deal with the scheduling problem, in that case online-scheduling reporting good results with a CPU time of less that 100 seconds for a 14 day schedule. Genetic Algorithms have also been applied to the Job-shop problem with reentrant product flows as shown in the work of Noze et al. (1999).

A novel optimization technique, the ant colony optimization (ACO) algorithm is also used in a recent work of Jayaraman et al. (2000). This algorithm is inspired in the way that have the ants to find the shortest path from their nests to the food source and back. They compare their results which those obtained by GA and also, in small problems, with the optimal solution found using mathematical programming. They report a 100% efficiency in finding the optimal solution for small problems and 92 % of success for larger problems.

The application of heuristics to the scheduling problem represents the current industrial practice. In most cases, because of the deep knowledge of the process it is possible to develop knowledge-based heuristics that allow to obtain good solutions to the scheduling problems of the industrial site taking into account specific issues that can be either difficult or impossible to be included in a general model. The application of Knowledge-Based scheduling systems has also been reported in the academic literature as the work of Sauer and Bruns (1997) who develop a system capable of representing the scheduling knowledge, allowing flexible reuse and adaptation of scheduling algorithms.

2.2.2 Mathematical programming

A major part of the research on scheduling applied to multipurpose plants contemplates the development of mathematical programming approaches. Shah (1998) makes an extensive review of the planning and scheduling problem in this respect.

According to the time domain representation the scheduling approaches can be classified into two types: discrete and continuous time representations. Discrete representation (see figure 2.2) is based on a uniform discretization of the time domain. One of the most relevant work based in the discrete representation of the time is the formulation by Kondili et al. (1988, 1993). Kondili introduced the representation of the scheduling problem as an State Task Network (STN). The STN representation uses two types of nodes: the *state* nodes, representing the feeds, intermediate and final products and the *task* nodes, representing the processing operations which transform material from one or more input states to one or more output states. State and task nodes are graphically represented as circles and rectangles respectively.

The STN representation simplifies the scheduling problem. One task can start if it has enough input materials (states) regardless of a possible recipe structure , as it does not use the concept of recipe. The MILP formulation presented by Kondili et al. (1988, 1993) is based on binary variables that indicate whether tasks start in specified pieces of equipment at the start of each time period. The number of batches is not fixed (in fact the concept of batches is nonexistent due the STN representation) and the objective function is to maximize the profit. The formulation is very complete as it also introduces sequence and frequence dependant

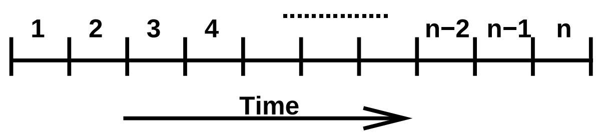

Figure 2.2: Discretization of time with fixed intervals

cleanings and management of intermediate storage. The main drawback of this formulation is the quantity of the necessary binary variables that causes long computational times.

The work of Kondili is the basis of other works that modify the original formulation in order to improve the numerical performance. Shah et al. (1993) that modified the allocation constraints in order to decrease the integrality gap and Papageorgiou and Pantelides (1996a,b) that contemplated the campaign planning/scheduling problem are some examples in this area of research.

Another interesting work is presented by Rodrigues et al. (2000). In this work time windows and constraint propagation is used as an effective way to solve the MILP formulation of Kondili et al. (1993). The time windows processing generates a set of new constraints which allows solving the initial model with very short computing times.

A new kind of representation called RTN (Resource Task Network) was presented by Pantelides (1994). This new representation was based on consuming and producing "resources" instead of materials. This new representation allows a simpler representation of the changeovers and resource management. The associated mathematical model presented in the same work was also based in a uniform representation of the time, like the previous works based on the STN representation.

In both approaches STN or RTN approach, the formulation of the MILP models using a uniform time discretization model have some inherent problems:

- The difference in the time of operation for each task involved in the schedule, may force to use a large number of intervals. Such increase in the number of intervals generates a bigger model that can be difficult to solve.
- The modeling of continuous and semicontinuous operations is difficult. As the time is discrete, the operation time for this kind of operations should be adapted to the time model used.
- Is difficult to model operations with operation time dependent on the batch size.

In order to avoid these kind of problems, several continuous time models have been developed. An example is the work presented by Pinto and Grossmann (1996) where a continuous time model was used. Time slots were defined according to pre-ordering constraints resulting on a model that reaches the optimal solution with short calculation times. Each unit has assigned a set of non-uniform slots as shown in figure 2.3. The starting and finishing time for each slot is calculated within the model.

Ierapetritou and Floudas (1998b,a); Ierapetritou et al. (1999) presented a complete scheduling model based on a continuous-time discretization and the STN representation. In these

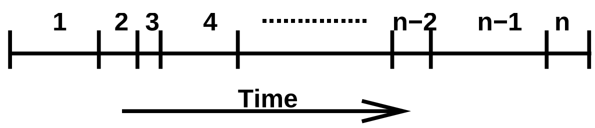

Figure 2.3: Discretization of time with variable intervals

works, task events are decoupled from unit events and time sequencing constraints are included. Only one event list is considered. Binary variables relate the events with the start of the tasks and the utilization of the unit. An iterative procedure is presented in order to find the minimum number of slots/event points that are needed to find the optimal solution. This formulation is later revised in a work of Castro et al. (2001). The revision discovered that the iterative procedure developed to find the optimal number of slots proposed by Ierapetritou didn't guarantee the optimum.

Mockus and Reklaitis (1999a,b) also developed a continuous time discretization model (fig. 2.3) based on a STN representation. The formulation presented in this paper leads to a MINLP model which can handle batch and continuous processes. Sequence dependent changeovers and multiple demands are also considered. The objective function is to minimize costs taken into account operating costs, sales revenue and inventory costs within a defined time horizon. A Bayesian heuristic is used to solve the large MINLP problem generated.

Mendez et al. (2001); Mendez and Cerdá (2000) also develop continuous time models. The main difference between the models presented in the last papers and the previous ones is that the use of sequence constraints as binary variables is introduced. In this way the use of an explicit time discretization is avoided resulting on smaller and easy to solve model.

2.3 The scheduling problem and the industry

The scheduling problem has been around for a long time. But it is not until the 1980's that the firsts commercial scheduling packages were available to the industry (Shobrys, 2001). Some of the first firms to offer scheduling products were OPT, i2 and Numetrix.

In the 1990's the market for scheduling products booms and a big number of new scheduling products began to appear. One of the characteristics of this new generation of products was that the capability to be connected to other systems is increased. In the mid 90's some vendors started to migrate their systems to Microsoft Windows NT platforms either via a client-server architecture or moving the entire application to the Windows NT. Also more intuitive interfaces were developed.

Sadowski (1998) estimated the revenues for Advanced Planning and Scheduling (APS) packages in the year 2000 in more than 3000 million EUR with a growth rate of 70 percent with nearly 100 different scheduling packages available (LaForge and Craighead, 2000).

Most of the APS available packages are designed with the job-shop scheduling problem in mind but according to Graham (2000) some of them also deal with some of the characteristic process industry problems like tanks, intermediate storage, change-overs and cleanings.

According Sadowski (1998) the different scheduling packages available in the market can be categorized as follows:

- *Job scheduling packages*: There are the simplest packages. Somehow they mimic the method that a scheduler might use to develop a schedule using a traditional magnetic board. They usually have a large number of scheduling rules available and are very fast methods. No optimization is performed and are not useful if special constraints exist (i.e. sequence-dependent setups)

- *Resource-based scheduling packages*: Are often based in the theory of constraints. The idea is to schedule first the jobs that need the bottlenecking resource and then allocate the remaining jobs to the non-bottleneck resources using a combination of backward and forward scheduling algorithms. These packages can produce good results if there are true bottlenecks in the system. The solution time is greatly dependent on the package.

- *Event scheduling packages* : Employ simulation concepts to develop a schedule. The start is at the current time trying to schedule any possible operations. There is a stepping forward in time until conditions change to allow additional operations to be scheduled. The method for selecting which operations should be scheduled first can be resource-dependent and dynamic (i.e. minimize setup time). This method tends to produce schedules with fewer gaps, but the solution time can be long. They produce good solutions to systems where sequence-dependent setups or other special constraints are present.

- *Optimization scheduling packages:* Some APS packages provide optimization methods, generally based on some type of search procedure, with the quality of the solution directly dependent on the amount of solution time allowed. Such methods tend to work best for very complicated scheduling problems but require a very long time to generate a good solution. They are not well suited to work with systems that have day-to-day changes.

According to a survey that took place in the United States (LaForge and Craighead, 2000) there is a big interest in implementing a scheduling solution in the industry. The top goals achieved by the industries with an scheduling system implemented were: on-time deliveries, high-quality products, good costumer relations, schedule reliability and quick response to changes. Almost half of them (48%) perform rescheduling at least once a day whilst more than one-third (34%) regenerate the schedule only once each week. Less than half of the plants have its scheduling system fully integrated with other plant information systems. The most important conclusion of that study is that plants reporting to have a bigger improvement in the production after the implantation of the scheduling system were the plants having the scheduling system integrated with upstream planning systems (ERP's) and downstream plant systems and that regenerate the schedule at least once per day.

The advantages of using a scheduling tool are also clear. Hess (2000) stated a typically average of 15% increase in the production, a 40% reduction in inventory levels and between 80% and 90% improvement in on-time delivery orders.

Some fonts (McCall, 2001; McGettrick et al., 2001) state that the implementation of a scheduling package generates a ROI (Return Of Investment) of 30% to 300% . In a case study applied to the pharmaceutical industry (Thomas, 1999), the ROI was even higher as the system paid for itself within three months of go-live. In the same paper, the author also explained the excellent figures of impact in the business that results on that specific implementation of a scheduling package (i.e. lead time reduced from sixteen weeks to two days). Morris (2000) revise the results obtained after the implantation of scheduling software in different food factories showing again the clear advantages obtained after the use of a scheduling tool.

A key issue for the successful application of a scheduling system is the integration with other plant levels. As shown (Stene, 1999) not all the scheduling packages available have connecting capabilities. Some of them are integrated in specific ERP vendors and are not available to exchange data with other products. In the other hand many of them are occasionally "open" thus being able to exchange data with other plant levels. Ferrar (2001b) also stated that many of the ERP vendors available offer a "pre-integrated" APS solutions that can help with the integration task.

In order to deal with the integration aspects, two standards have been deployed: the ISA SP88 and ISA S95 (International Soc. for Measurement and Control, 1995, 1999, 2000).

As explained by Haxthausen (1995) the ISA standard S88 is focused in batch control. The standard is split in two parts:

The first part is focused on the models and terminology of batch control. The definitions in that standard allow the clear verbal exchange of batch control concepts between different people. The model supplies a hierarchical definition of the recipe and the equipment information .

The second part of the S88 standard defines a common database for data exchange between different software that works with batch recipes. The objective is to make it easy for the different vendor systems used in batch manufacturing to interoperate. The data model in S88.02 provides a standard reference for the objects and attributes found in batch systems. This allows interface specifications to define information-exchange methods, and to ensure that the data can be understood by different vendor systems. Actually several batch control systems incorporates the S88 guidelines.

Complementary to the S88 standard a new standard, the S95.01 (International Soc. for Measurement and Control, 2000) emerged focused in the integration into larger business systems. As explained by Nowicki and Brandl (2000) the objective of S95.01 is to reduce the need for custom integration solutions, simplify multi-vendor integration, and improve reusability and transportability of functions across the enterprise. The base model of S95.01 provides a hierarchical view of the activities associated with manufacturing enterprises. This model identifies specific functions in an enterprise and defines how these functions interact with the production control functions of manufacturing.

These standards are focused on providing a common language that eventually allows integration between the different information levels available in the plant. As shown by Shobrys and White (2000) the integration of the different levels of the enterprise: planning, scheduling and control can improve the production process reducing working capital (due to the reduction of stocks of finished products and raw materials) by a factor near to 50 %.

2.4 Conclusions

The scheduling problem has been a common topic of research for a long time. Despite the large number of works related to this topic, there is still a lack of a solution that can cope with the complexity of the problem as exposed in the previous chapter.

Different approaches have been developed in order to find a solution of the scheduling problem. Relevant works related to heuristic and mathematical programming approaches have been presented in this chapter.

In a last part of this review, several industrial aspects have been presented. The benefits of applying a scheduling tool in the industry are clear as has been shown by different fonts. High ROI’s and important improvement figures are the key aspects that shows how important is the research in the scheduling an planning area for the industry.