UNIFIED FRAMEWORKS
FOR
OPTIMAL PROCESS PLANNING AND SCHEDULING

Constantinos C. Pantelides
Centre for Process Systems Engineering
Imperial College of Science, Technology and Medicine
London SW7 2BY
United Kingdom

Abstract

This paper is concerned with unified frameworks for the description and solution of a variety of process scheduling problems, including various types of production scheduling for single plants, maintenance scheduling, and multisite plant scheduling.

Two such unified frameworks are presented, based on representing processes as State-Task Networks and Resource-Task Networks respectively. The former leads to a number of related mathematical formulations for various process scheduling problems, while the latter allows the development of a single simple mathematical formulation based on the uniform characterisation of all available resources.

Keywords

Flexible Plants, Planning, Scheduling, Process Optimisation.

Introduction

The operation of flexible plants involves the satisfaction of a number of production requirements placing competing demands on a set of limited resources, such as processing equipment, storage capacity, utilities, and manpower. The problem of efficient resource utilisation leads to a class of scheduling problems which have received considerable attention over the past couple of decades. A detailed taxonomy of these problems based on the types of resources taken into account and the operating mode of the plant under consideration has emerged. The very significant body of literature in this area has recently been reviewed by Reklaitis (1992) and Rippin (1993).

Scope of Process Scheduling

Much of the research effort to date has focussed on the scheduling of production for individual plants situated at a single geographical site and involving a set of batch, semi-continuous or even continuous unit operations. As is well known, this is in itself a complex problem, optimal or even feasible solutions to which are often notoriously difficult to obtain. However, it must also be recognised that production scheduling is only one aspect of the wider problem of process scheduling. For instance, the scheduling of plant maintenance operations, the co-ordinated planning of the production at a number of distinct geographical locations, and the management of distribution and supply chains, all lead to important scheduling problems that interact strongly with production scheduling at individual plants.

Even at the level of single plant operation, recent technological advances in plant automation now permit, at least in principle, the implementation of much more complex operating strategies than has been practicable in the past. This development, in turn, opens the way for the abolition of artificial restrictions, often introduced with the aim of simplifying plant operation. In the case of batch plants, one such restriction is maintaining integrity of individual batches throughout the plant when this is not dictated by overriding safety and/or regulatory considerations. The value of well-established concepts in the parallel utilisation of multiple equipment items, such as the ideas of "in-phase" and "out-of-phase" operation, also comes under question. Of course, any additional flexibility almost always results in an increase of the complexity of the underlying production scheduling problem.

Moreover, the operation of flexible plants involving energy-integrated operations (*e.g.*, Corominas *et al.*, 1993), and the emergence of new types of plant, such as the "pipeless" plants recently reported in the literature (Niwa, 1991), have led to different types of production scheduling problems.

Finally, the strong interactions between process scheduling on one hand, and the design of flexible plants on the other, are now well recognised. Thus detailed scheduling considerations are taken into account at the plant design stage (Birewar and Grossmann, 1990).

The Need for Unified Frameworks for Process Scheduling

It could be argued that the complexity and diversity of process scheduling problems naturally give rise to the need for an equally diverse set of methodologies and techniques for their solution. Indeed, to a large extent, this has been the approach followed in much of the literature.

However, even for the problem of production scheduling considered in isolation, realistic industrial applications rarely fall neatly within the confines of a single type of scheduling problem. In our experience, for instance, the common classification of such problems as "equipment dominated" with or without intermediate storage capacity, and "resource (*i.e.* utility) constrained" (see Reklaitis, 1992) is very often of rather limited usefulness: most real plants are subject to simultaneous restrictions in processing equipment, intermediate storage, and utilities. Furthermore, when the provision of these resources in a plant is finely balanced, small changes in the input data (*e.g.* production requirements) often cause the bottleneck to shift from one resource to another.

The above factors, together with the strong interactions between different types of process scheduling problems, seem to advocate the need for the establishment of general frameworks for the representation and solution of such problems, thus ultimately opening the way for the development of truly general-purpose scheduling algorithms and software.

This paper presents and discusses two different unified frameworks for process scheduling. The first one, based on the *State-Task Network* (STN) process representation originated as a mathematical formulation (Kondili *et al.*, 1988, 1993), and solution algorithm (Shah *et al.*, 1993a), for production scheduling problems over short planning horizons. The overall approach has since been extended to other types of production scheduling problems for conventional batch and semi-continuous plants (Shah *et al.*, 1993b, Crooks *et al.*, 1993). It has also been applied to pipeless plants (Pantelides *et al.*, 1992), and plants involving heat-integrated unit operations (Papageorgiou *et al.*, 1993a).

The application of the STN framework to diverse process scheduling problems leads to a number of different, yet conceptually related, mathematical formulations. The second unified framework presented in this paper is based on the representation of processes as *Resource-Task Networks* (RTNs), in which abstract tasks transform one set of production resources into another, with all resources being handled in a uniform manner. It is shown that this can lead to a *single* simple mathematical formulation describing a very wide variety of production and other process scheduling problems.

The rest of this paper is organised as follows. The next section reviews the STN process representation, already presented in detail elsewhere in the literature, and discusses the basic assumptions made with regard to plant operation and production resources. It then considers a number of different process scheduling problems, all of which can be accommodated within the STN framework. It finishes with a critique of the overall approach, identifying its strengths and weaknesses. The following section discusses the key ideas and concepts underlying the RTN process representation, and demonstrates how these address the weaknesses of the STN framework. It then presents a simple derivation of a unified mathematical formulation for process scheduling. Some concluding remarks are made in the final section.

The State-Task Network Unified Framework

Process Representation

State-Task networks (Kondili *et al.*, 1988, 1993) represent process recipes as transformations of material ("states") effected through a set of processing steps ("tasks"). A typical STN is shown in Fig. 1, with tasks represented as rectangles, and states as circles.

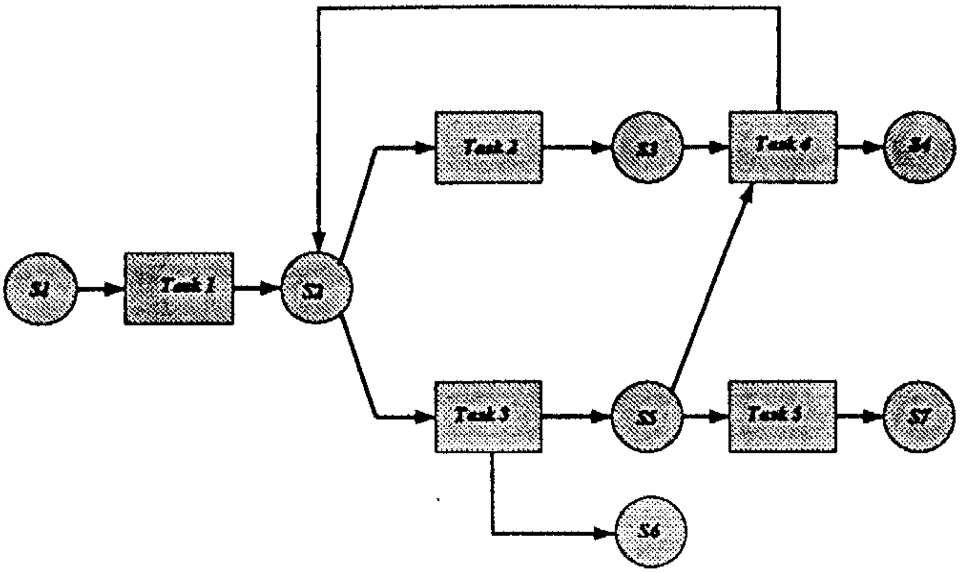


Figure 1. State-Task Network Process Representation

The STN representation allows the unambiguous description of recipes involving shared intermediates, multiple processing routes to the same intermediate or product, and material recycles. In order to limit the complexity of the underlying mathematical problem, a particularly simple model is adopted for the processing tasks. Thus it is assumed that tasks consume their input states in fixed, given proportions, and produce their output states also in fixed given proportions. The processing times associated with the various output states of a task are also fixed and known *a priori*, but may be different for different states.

Production Resources

The STN is a representation of the process, and does not contain any information regarding the available resources. However, the latter are associated with either the tasks or the states in the network. Thus, each item of *processing equipment* is assumed to be suitable for a subset of the tasks in the STN, with a given minimum and maximum capacity for each such task. Furthermore, equipment may require cleaning or some other preparation between being used repeatedly for different tasks — or even the same one. The need for cleaning operations may be determined either by the precise sequence of processing tasks performed in the vessel, or alternatively by the frequency of utilisation of the vessel. Taking account of such operations is important, not only because they prevent the vessel from carrying out production tasks, but also because they may themselves pose demands on utilities, manpower, and cleaning materials.

Each task may also employ a subset of the available *utilities* (*e.g.* steam, cooling water, manpower *etc.*). The amount of each utility consumed may vary over the duration of the task; the consumption at any given time during the execution of the task may be fixed or variable (*i.e.* depending on the batchsize), or a combination of the two.

Given amounts of *dedicated storage capacity* are assumed to be associated with the various states representing the raw materials, the intermediates and the end-products in the process.

A number of features of flexible plant operation can be accommodated simply by modifying the basic process STN. For instance, *resource substitution, i.e.* the ability of a task to be carried out using two or more alternative subsets of utilities, is easily handled by splitting the task into two or more tasks that have the same input and output states as the original task but make use of different utilities — and possibly differ in other ways (*e.g.* with respect to their processing times). Also, *multipurpose storage capacity, i.e.* the ability of a storage vessel to be used for storing material in different states at different times can be modelled by introducing special material "holding" tasks operating on the relevant states, and modelling the vessel as a processing equipment item that is suitable for all or a subset of these new tasks. The use of processing equipment as *temporary storage* capacity for material produced in them following the completion of a task can also be modelled by the introduction of such holding tasks.

Fig. 2 illustrates a number of such modifications applied to the basic STN of Fig. 1. In particular, the introduction of a holding task on state $S3$ allows the use of processing equipment employed for task 2 as temporary storage for this material. Similarly, splitting task 3 of Fig. 1 into two tasks (denoted as 3 and 3′ in Fig. 2) allows the use of alternative sets of utilities by the two tasks while still carrying out the same transformation of materials.

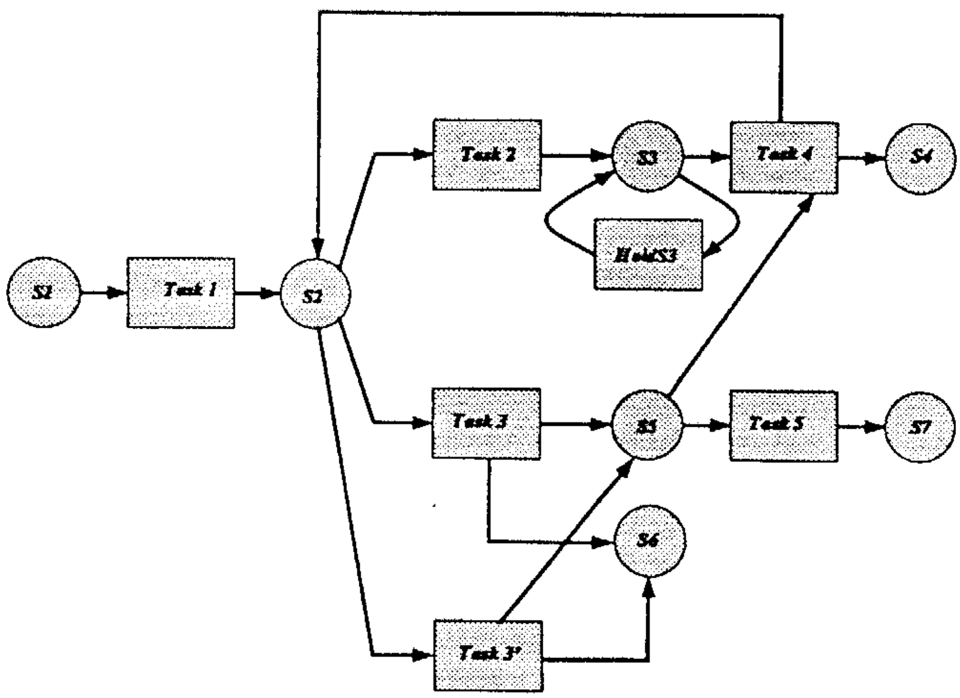


Figure 2. Modelling Resource Substitution and Multipurpose Storage Capacity

Short-term Production Scheduling

The problem of short-term production scheduling aims to establish the optimal manner of utilising the plant resources over a given time horizon, possibly satisfying a number of production requirements for different products at different times during and/or at the end of the horizon. The availability of processing equipment, and the level of availability and unit cost of the various utilities over the horizon are assumed to be known, as are the initial amounts of the various materials held in storage. The objective function to be maximised is typically an economic performance criterion taking into account the value of the products, the cost of raw materials and utilities, and any penalties incurred because of failure to satisfy minimum production requirements within given nominal deadlines.

Short-term scheduling is appropriate for flexible plants, the production of which is geared primarily towards satisfying individual customer orders. This leads to relatively short time horizons, typically of the order of one day to a few weeks. Due to the diversity of the orders that have to be fulfilled over successive planning horizons, and the small production amounts typically involved for individual products, no regular production pattern is established[1].

A fundamental issue in the mathematical formulation of any scheduling problem is that of the representation of time. Kondili *et al.* (1988, 1993) based their formulation on a discretisation of the time horizon into a number of time intervals of equal and fixed duration. All system events (such as the start and the end of task processing, and changes in resource availability and/or unit costs) are forced to occur at the interval boundaries. This provides a simple mechanism for ensuring that resource limitations are satisfied at all times by ensuring that they are satisfied at a finite number of time grid points. On the other hand, obtaining optimal solutions with this approach may require the use of fine grids, thus leading to large optimisation problems.

Kondili *et al.* (1993) showed that it is possible to formulate the general short-term production scheduling problem taking into account a wide variety of the problem features as a mixed integer linear programme (MILP). The key variables in the formulation relate to the utilisation of the various resources over the time horizon. Specifically they characterise (a) the task for which each item of processing equipment is being used over any given time interval if it is not idle, and the associated batchsize, (b) the amount of material of each state held in dedicated storage over each interval, and (c) the amount of each utility being used over each interval.

For realistic applications, the MILP resulting from the above formulation may involve several thousands of binary decision variables. Shah *et al.* (1993a) presented a number of measures aiming both at reducing the integrality gap of the mathematical formulation, and at accelerating the solution procedure using a modified branch-and-bound method. Overall, these techniques allow relatively large problems to be solved with reasonable computational

[1] Such plants are sometimes referred to in the literature as operating in a "short production campaign" mode. However, as observed by Reklaitis (1992), no real campaign structure exists in this case, and analogies with campaign mode operation are not particularly useful.

effort.

One of the deficiencies of the original Kondili *et al.* formulation was the way in which sequence-dependent cleaning requirements were expressed mathematically, with the number of necessary constraints being of $O(H^2)$, where H is the number of discretisation intervals over the entire time horizon of interest. This problem was addressed by the later work of Crooks *et al.* (1993) who introduced the notion of the *unit State-Task Network* (USTN). This is a network representation of the different states in which an item of processing equipment can exist (*e.g.* "clean", "dirty", "cold", "hot" *etc.*), and the way in which its state changes when the vessel is used to carry out a task. Only allowable state transitions are included in the network. For instance, Fig. 3 shows the USTN describing the operation of a vessel that may be used for the manufacture of either a dark or a light dye. At any given time, the vessel can exist in one of three states: clean, dark-soiled, and light-soiled. As can be seen, a vessel in the light-soiled condition may be used to produce a dark dye without any cleaning, but the opposite is not true.

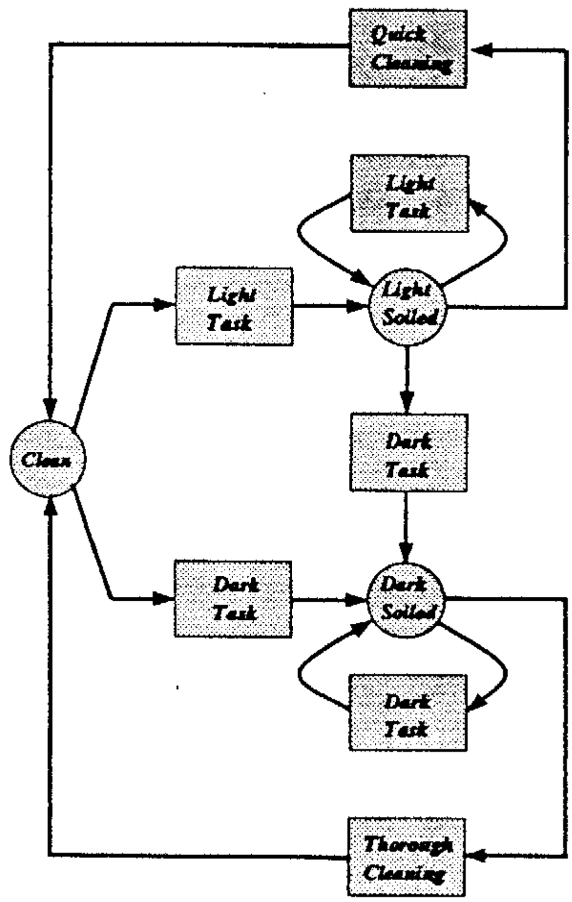


Figure 3. Example of Unit State Task Network

Crooks *et al.* introduced a new set of variables characterising the state of a vessel over the time horizon, and a set of constraints that ensure that changes in this state are correctly accounted for. This allows the modelling not only of sequence-dependent cleaning requirements, but also of more general task pre-conditions for unit utilisation, such as, for instance, the need for heating up a reactor vessel before the introduction of the reactants.

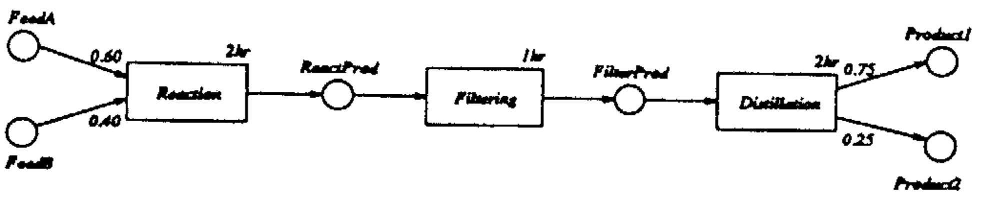
The exploitation of energy integration in the operation and design of flexible plants is a relatively new area of research (see, *e.g.*, Vaselenak *et al.*, 1986; Corominas *et al.*, 1993). In practice, because of the relatively high value of the products produced in many multipurpose plants, and the paramount importance of other factors (such as maintaining high levels of customer service through the timely satisfaction of orders), energy conservation is rarely a primary scheduling objective. Furthermore, the energy that can be recovered from many batch operations is often of too low a grade to be useful. Nevertheless, potential for energy integration *does* exist in some cases, especially in such areas as brewing and other biotechnological processes.

The use of heat integrated operations interacts strongly with the production scheduling as a certain degree of synchronisation with respect to time must be introduced between the tasks producing energy and those consuming it. Papageorgiou *et al.* (1993a) demonstrate that, in fact, this can be achieved relatively easily within the context of the STN framework. They consider both *direct* energy integration in which the operations exchanging heat must overlap in time, and *indirect* integration in which a heat transfer medium (HTM) may be used to store energy over limited periods of time. In either case, there is normally the additional flexibility of running each of the heat-integrated operations independently, relying wholly on external utilities. By introducing the cost of such utilities in the scheduling objective function, the overall most economical solution that satisfies the production requirements and other constraints can be determined.

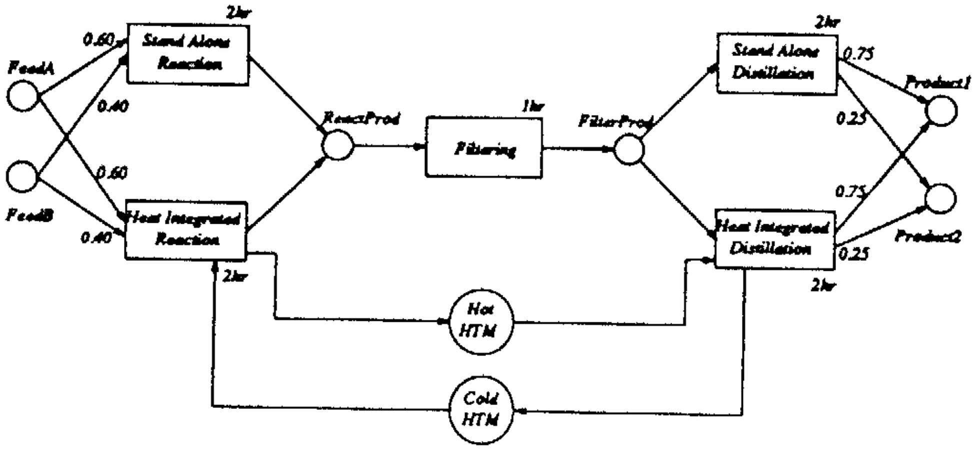
In the indirect energy integration case, the heat transfer fluid at different temperatures is treated as different states of material in the STN. This is illustrated by the simple example of Fig. 4. The original STN (Fig. 4a) involves a batch exothermic reaction forming a liquid mixture and a solid sediment. Following filtration to remove the sediment, the liquid mixture is separated in a batch distillation step. Potential for heat integration exists between the reaction and the distillation tasks, and this is to be realised through a heat transfer medium. However, the two tasks can also take place in a "stand-alone" fashion, if necessary. The required STN modification is shown in Fig. 4b. Additionally, a new class of constraints may have to be introduced in the scheduling formulation to describe the limited storage lifetime of hot HTM.

Periodic Production Scheduling

The short-term production scheduling problem considered above is relevant to plants driven primarily by the need to satisfy individual customer orders, with little similarity in either product demands or plant operation from one planning period to the next. In contrast, periodic scheduling is more appropriate for plants operating under more stable demand conditions over relatively long planning horizons, typically ranging from one month to a year. This allows considerable simplification of plant operation which now involves the same sequence of tasks executed repeatedly at a fixed frequency. A periodic (cyclic) operating mode is thus established.



(a) Original State-Task Network



(b) Modified State-Task Network

Figure 4. STN Modifications for Indirect Energy Integration

The periodic production scheduling problem has received extensive attention in the process engineering literature, mainly in the context of campaign-mode plant operation (Mauderli and Rippin, 1979; Wellons and Reklaitis, 1989a,b, 1991a,b). The latter involves the division of the planning horizon into a number of production campaigns, with all available plant resources being dedicated to the production of a small subset (possibly one) of the plant final products and stable intermediates. As campaigns are relatively long, a periodic operating mode is established over each one of them. Thus the associated production planning problem involves the determination of both the campaign structure (*i.e.* the number, duration and product mix of the campaigns), and a detailed operating periodic schedule for each.

A key concept in periodic production scheduling is that of the *cycle time*. For plants producing identical batches of a product, this can be defined as the average elapsed time between the production of successive batches. Alternatively, if the size of different batches of the same product varies depending on the path followed in the plant, then a "cycle" may be defined in terms of the production of exactly one batch along each of the paths considered, with the cycle time being the shortest possible time between the start of successive cycles (Wellons and Reklaitis, 1989a). If more than one product is produced simultaneously in the plant, a different cycle time is defined for each.

The extension of the STN framework to the periodic scheduling of multipurpose plants was considered by Shah *et al.* (1993b). They note that the above cycle time definitions are not always satisfactory for general multipurpose plants in which different products may be coupled through shared intermediates, and batch identity is not necessarily preserved. Instead, they observe that the cycle time is directly related to the complexity of the periodic operating pattern for a given plant. Longer cycle times simply allow more complex sequences of operations to be carried out. In fact, it could be argued that, in general, the "optimal" operating schedule for a plant from a narrow economic point of view is one which is not periodic at all. However, such "optimality" has to be balanced against the underlying complexity of operation and also the ability of currently available algorithms to derive optimal non-periodic schedules for very long time horizons. Consequently, the acceptable level of operating complexity provides an upper bound on the length of each cycle, but the latter is otherwise free to be determined together with the detailed set of operations to be performed.

As demonstrated by Shah *et al.* (1993b), the mathematical formulation of the short-term scheduling problem can be extended to the periodic scheduling problem in a relatively straightforward fashion. They show how the problem can be formulated in its entirety by focussing on operations over a single cycle of a given duration despite the existence of tasks that span more than one consecutive cycle. This is achieved by considering tasks extending beyond the end of one cycle as notionally "wrapping around" to the start of the cycle. The optimal cycle duration can be obtained by iterating between given lower and upper limits, solving a sequence of fixed cycle time problems. An alternative single-level formulation treating cycle time as an unknown determined simultaneously with the optimal periodic schedule has been presented by Shah (1992).

An interesting result shown by Shah *et al.* (1993b) is that, even for very simple linear STNs involving as few as three processing tasks leading to a single product, the optimal schedules violate many of the largely artificial operating restrictions imposed on the utilisation of processing equipment by much of the earlier literature. Thus, in some cases it is advantageous to use parallel equipment both in-phase and out-of-phase over the same cycle. In other cases, the optimal solution involves using the same item of equipment for more than one different task during a cycle even when this gives rise to some sequence-dependent cleaning requirements.

Campaign Planning and Scheduling

As already noted, periodic production scheduling occurs as an important sub-problem within the context of the overall campaign planning and scheduling problem. Earlier approaches (Mauderli and Rippin, 1979; Wellons and Reklaitis, 1989a,b, 1991a,b) advocated a hierarchical approach in which a set of dominant campaigns was identified, each involving a complete periodic schedule for the production of a subset of the products and stable intermediates of the plant. A campaign planning step is then undertaken to allocate the available production time among the various campaigns.

A large variety of campaign planning and scheduling problems may be solved using the hierarchical approaches described above. However, as observed by Reklaitis (1992), these suffer from a number of limitations when applied to general multipurpose plants. The key restriction arises from the need to identify all dominant campaigns *a priori* irrespective of which campaigns will actually be used by the final plan. Reklaitis (1992) notes that, for a plant with P products, the identification of the dominant campaigns could involve the solution of 2^P separate multiple campaign problems, and that the number is even greater for cases involving utilities that can, in principle, be used for multiple products. In the latter case, one would have to decide *a priori* the distribution of the utility among the various products in the campaign, thus leading to even more distinct combinations to be considered.

In fact, a more fundamental problem arises when the same resource is allowed to be used for more than one product at different times within the *same* campaign. In such circumstances, the possible distributions of the resource are difficult to enumerate, let alone assess. In fact, even the formulation of an appropriate objective function for the periodic scheduling problem in any single campaign is in itself problematic. This is due to the interactions between the various products involved, and the complex relationship between the objective of the single campaign and that of the overall campaign planning problem.

In view of the above problems, Papageorgiou and Pantelides (1993) proposed a simple hierarchical procedure that aims to improve campaign plans derived using other algorithms. The approach followed is to take each campaign in the production plan and apply to it the optimal periodic scheduling algorithm of Shah *et al.* (1993b). Because of the additional flexibility in the utilisation of resources allowed by the latter, this may well lead to improved throughput for individual campaigns. A final campaign re-timing step aiming to extract maximum advantage of the increased throughput is then undertaken. This is similar to the campaign planning step employed by other hierarchical algorithms.

Of course, approaches such as the above cannot be guaranteed to obtain optimal schedules, although they do tend to improve existing ones in many cases. In principle, true optimality can be assured by considering the entire campaign planning and scheduling problem in a single-level mathematical formulation that simultaneously determines both the campaign structure and the optimal periodic schedule for each campaign. This approach has several advantages. First, it automatically determines how the production of a product may be subdivided into the production of several stable intermediates taking place in different campaigns[2]. Secondly, the optimal campaigns are formed without the need for selecting them out of a large set of *a priori* identified candidates. Thirdly, only an overall economic criterion for the entire production (similar to that used at the campaign planning step of hierarchical approaches) is utilised, without the need for formulating artificial objectives for individual campaigns. Finally, the incorporation of general periodic scheduling formulations allows extensive flexibility in processing structure and resource utilisation. Of course, for problems involving large number of products and/or many campaigns, the size of the resulting mathematical programming problem may be prohibitively large, and the use of mathematical decomposition techniques may be inevitable.

[2] Reklaitis (1992) identifies the failure of existing algorithms to do this as one of their key limitations.

Maintenance Scheduling

The problem of maintenance scheduling is concerned with the determination of the timing of one or more maintenance activities to be carried out on a given plant. These may include both preventive maintenance operations performed on equipment that is still functioning properly, and operations involving the repair of process equipment that is currently inactive because of an earlier malfunction.

Maintenance planning in itself often leads to interesting scheduling problems aiming to determine the best utilisation of limited resources such as maintenance crews. Moreover, there are invariably strong interactions between maintenance and production planning, as preventive maintenance temporarily takes functioning processing equipment away from production, while repairs are a necessary prerequisite for equipment to be returned to normal production activities.

Dedopoulos and Shah (1993) have recently extended the basic STN framework to cover combined short-term scheduling of maintenance and production in multipurpose plants. They consider a set of maintenance crews with a variety of skills, and a set of maintenance jobs that may potentially be carried out on the plant equipment over the time horizon of interest. Detailed factors such as limited duty periods (*i.e.* the length of time that a crew may operate without interruption), statutory rest periods, crew call-out charges, job precedence relations *etc.* are also taken into account. The resulting mathematical formulation of the problem as a MILP is an extension of that of Kondili *et al.* (1993) incorporating additional variables to characterise crew utilisation and the timing of the maintenance tasks.

Other Process Scheduling Problems

Pipeless plants have recently been described in the literature (see, for instance, Takahashi and Fujii, 1990; Niwa, 1991; Shimatami and Okuda, 1992; Zanetti, 1992) as alternatives to more conventional batch plants. Their key characteristic is that, instead of having complex piping interconnections between processing equipment, they use a set of transferable vessels to move material from one processing station to another. The vessels may be fitted with their own locomotion mechanism, or they may rely on a shared pool of automated guided vehicles (AGVs) for their movements.

Pipeless plants are useful for processes involving solids or slurries that are difficult to transport in pipes. They are also particularly advantageous for processes with extensive cleaning requirements: most of the cleaning in pipeless plants is associated with the transferable vessels and can therefore be carried out at specialised cleaning stations, thus minimising the effect on the availability and productivity of processing stations.

Production scheduling for pipeless plants is similar to that for conventional plants except that the utilisation of some additional production resources has to be characterised. These include the transferable vessels and, where appropriate, the AGVs. Pantelides *et al.* (1992) extended the concept of the STN to pipeless plants by using different states to model different types of material temporarily stored in transferable vessels as well as material stored in dedicated storage vessels. Explicit charging and discharging tasks corresponding to the transfer of material between the two types of vessels are also introduced. Empty vessels, either soiled or clean, are modelled as states, and cleaning tasks are used to transform the former into the latter. The resulting MILP formulation includes additional variables describing the movement of vessels (either loaded or empty) from one location to another, as well as the numbers of vessels waiting at various allowable locations in the plant.

Multisite production scheduling involves the co-ordinated operation of a number of plants situated at distinct geographical locations. Typically the plants are related through shared raw materials or common products. In some cases, one or more products of one plant may be used as raw materials for another. In its wider form, the scheduling problem also includes the management of the acquisition of raw materials from suppliers, and the distribution of the end products to customers.

In any case, the often significant geographical distances between the different sites of interest render the characterisation of transfers of material particularly important in this type of scheduling problem. Specifically in the context of the STN framework, the same type of material existing at two different locations can be modelled as two different states, possibly linked by one or more different transfer tasks. The latter typically share limited amounts of transfer resources, such as lorries, whose location and movement during the time horizon of interest must also be characterised. Lorries moving from one location to another may carry mixed loads involving material of more than one type, although some load combinations may be prohibited. Similarly, large warehousing facilities at the various locations may be shared among a number of different types of material.

Clearly, the material transfer issues described above also interact with the production taking place at individual sites. In principle, it is possible to model the detailed operation of both the production plants and the transfer network in terms of a single STN. However, obtaining operating schedules using the kind of techniques already described is probably both difficult (due to the size of the mathematical problem) and undesirable (due to the extreme centralisation of decision making that such an activity would imply). Thus a major challenge in this area is the development of appropriate ***aggregate*** production models for individual plants, which, when incorporated within the model of the entire network, always guarantee the feasibility of any derived schedule.

A Critique of the STN Framework

The STN framework clearly covers a very wide variety of process scheduling problems, leading to the development of a set of related mathematical formulations which are able to take detailed account of many of the features of industrial applications. The use of discrete representations of time certainly gives rise to large mixed integer optimisation problems involving thousands of binary decision variables. However, with careful formulation of the constraints and exploitation of the underlying mathematical structure (Shah *et al.*, 1993a), it has been possible to solve realistic problems, a number of which have been reported in the literature. Areas of application to date have included industrial additives (Shah *et al.*, 1993a), paints (Rapacoulias *et al.*, 1991), lubricants (Shah and Pantelides, 1991), polymers (Shah *et al.*, 1993b), dairy products (Crooks *et al.*, 1992) and others involving pharmaceutical, biotechnological and nuclear processes, and detergent manufacture.

The generality of the STN process description makes it a good basis for the development of general process scheduling software. In particular, Papageorgiou *et al.* (1992, 1993b) describe a prototype software system for the specification and solution of a variety of process scheduling and design problems for multipurpose plants.

However, the STN model of plant operation is rather restrictive in some cases. Thus, a task is always assumed to involve a transformation of material from one set of states to another, and to take place in a single item of processing equipment, with the possible use of utilities. Whilst this model is adequate for simple production scheduling applications, in more general cases, some tasks (such as cleaning operations in conventional or pipeless plants) are not necessarily associated with states of material, while others (such as movements of transferable vessels in pipeless plants, or lorries between plants) do not involve items of processing equipment. On the other hand, some tasks (such as heat-integrated operations) may involve more than one processing equipment item. Dealing with such complications requires various extensions of the basic STN concepts, as already described in this paper.

Another feature of the STN approach is that each item of processing equipment is treated as a distinct entity. However, characterising the utilisation of each item separately over the time horizon of interest tends to be inefficient in plants involving many identical items of equipment. In such cases, a more aggregate description of the processing equipment resource, similar to that used for other resources (*e.g.* utilities) would be more desirable.

Overall, the STN framework treats the various resources (such as material, processing equipment, storage capacity, utilities, and material transfer and transport devices) in a non-uniform manner. This non-uniformity has several important consequences. First, a relatively large number of different classes of constraints must be introduced in order to describe resource utilisation (*cf.* the basic formulation of Kondili *et al.*, 1988, 1993). Secondly, taking account of the features of novel types of process scheduling problems often necessitates the formulation of specialised constraints (*cf.* Pantelides *et al.*, 1992). Thirdly, establishing alternative or extended mathematical formulations (*e.g.* employing more detailed models of the operation of individual processing tasks, or describing uncertainty in resource availability) is rendered problematic by the need to take account of many different special cases.

The Resource-Task Network (RTN) framework described in the following section represents a systematic attempt to address the above issues.

The Resource-Task Network Unified Framework

Resources and Tasks

In his recent review of process scheduling, Reklaitis (1992) employs the term "resources" to include raw materials, utilities and manpower, but to exclude processing and storage equipment, pipe connections and transfer units. He further classifies resources into "renewable" and "non-renewable" types, depending on whether or not their availability is restored to the original level after usage by a task.

The main characteristic of the RTN framework is the entirely uniform description and characterisation of the available resources, with no distinction between equipment of any type and other resources. Furthermore, all resources are allowed to be produced as well as be consumed by tasks at any time during their execution. This has a number of implications:

- all types of material in the process (and not just raw materials) can be treated as resources;
- processing equipment items are treated as if they were "consumed" by processing tasks at the start of processing and "produced" at the end;
- some tasks may result in the generation of utilities (such as hot water or steam), either as the main output or as a by-product of their operation.

It should be noted that Reklaitis' renewable resources can be viewed as a special case of the general resources considered here, with a certain amount being consumed at the start of the task, and an equal amount being produced at the end.

The two fundamental concepts in the framework being established here are those of resources and tasks. A task is an abstract operation that consumes and/or produces a *specific* set of resources. Thus, in contrast to the STN framework, use of two alternative sets of resources (*e.g.*, a 10*te* reactor vessel instead of a 15*te* one) is always represented by two distinct tasks even if the same transformation of material is involved. The interactions between tasks and resources leads to the idea of the Resource-Task Network, a bipartite directed graph of the form illustrated in Fig. 5.

For a given plant, it is easy to identify the resources available, and the way in which they can be used for carrying out processing tasks. A more fundamental question is what precisely constitutes a distinct *resource type*. The key to classifying the available resources into the smallest possible number of distinct types is functional equivalence, which, in turn, depends on the detail of modelling employed. For instance, in simple production scheduling problems, two items of processing equipment belong to the same resource type if they have the same capacity, and are suitable for exactly the same set of tasks. On the other hand, if detailed consideration of plant connectivity is important (see, *e.g.* Crooks *et al.*, 1993), the two items will only be considered as parts of the same resource type if they have exactly the same connectivity. And if sequence-dependent cleaning requirements are important, the two items belong to the same resource only if and when they are in the same condition (*e.g.*, "clean" or "soiled"). Thus, the set of attributes which do or do not characterise a resource type is context dependent.

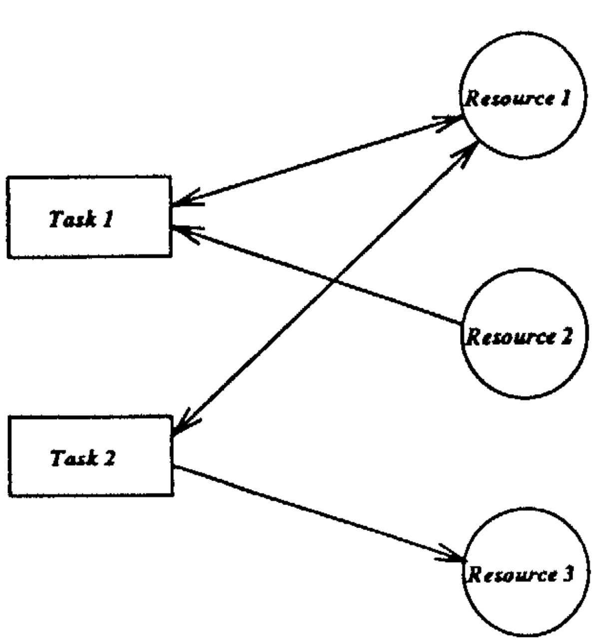


Figure 5. Resource-Task Network Process Representation

A further illustration of the issues arising in the categorisation of resources into resource types is provided by two amounts of material being held in two separate storage vessels. In the context of the STN framework, these would normally belong to the same "state" if their thermodynamic state is identical — although this may be unnecessarily restrictive if the relevant tasks in the process can tolerate small variations in composition and temperature. On the other hand, even if the thermodynamic state of the two amounts is identical, location may be an important distinguishing attribute in some applications such as multisite plant scheduling.

Task Extents and Resource Utilisation

The use of the RTN process representation permits the development of simple and elegant mathematical formulations. Furthermore, because of the general characterisation of resources, a *single* RTN formulation can describe a range of process scheduling problems. We demonstrate this point by presenting one such simple formulation based on a uniform, discrete representation of time.

We assume that each task k has a fixed duration τ_k. Resources are produced and consumed at a finite set of discrete times during the execution of the task.

Two variables, one integer and one continuous, are introduced to characterise the operation of a task k starting at time t. We denote these variables by N_{kt} and ξ_{kt} respectively. The *extent* of the task is defined as the pair (N_{kt}, ξ_{kt}) and determines the demands that the task places on the various resources. Specifically, the amount of resource r produced at time θ relative to the start of task k at time t is assumed to be related to the extent through a relationship of the form

$$\mu_{kr\theta} N_{k,t-\theta} + \nu_{kr\theta} \xi_{k,t-\theta} \tag{1}$$

where $\mu_{kr\theta}$, $\nu_{kr\theta}$, $\theta = 0, .., \tau_k$ are known constants. Negative values for these constants indicate consumption, rather than production, of the resource.

Consider, for instance, an exothermic batch reaction task k with a duration of 5 hours converting material A to material B. The reaction can be carried out in $10te$ glass-lined reactor vessels. It uses 0.25 kg/s of steam per tonne of material being processed during the first hour of its operation. After this initial period, the reaction reaches its normal operating temperature, and is kept there for the rest of the processing time by external cooling requiring 2 kg/s of cooling water per tonne of material being processed.

From the viewpoint of the RTN framework, it is clear that the task interacts with five distinct resource types, namely materials A and B, $10te$ glass-lined reactors, steam and cooling water. At the start of its operation, the reaction can be viewed as "consuming"[3] an amount of A, one or more reactor vessels, and an amount of steam utility. One hour later, the task "produces"[4] an amount of steam utility and "consumes" an amount of cooling water utility. Finally, at the end of its operation, the task "produces" an amount of material B, a number of reactor vessels, and an amount of cooling water utility.

In this case, we can interpret the extent variables N_{kt} and ξ_{kt} as the number of reaction batches starting in different reactor vessels simultaneously at time t, and their combined batchsize respectively. Assuming a time discretisation interval of 1 hour, the constants characterising this task are:

- Task duration $\tau_k = 5$;
- Resource $r = 1$ (Material A)
  $\mu_{k1\theta} = 0, \theta = 0, \ldots, 5; \nu_{k10} = -1.0, \nu_{k1\theta} = 0, \theta = 1, \ldots, 5;$
- Resource $r = 2$ (Material B)
  $\mu_{k2\theta} = 0, \theta = 0, \ldots, 5; \nu_{k2\theta} = 0, \theta = 0, \ldots, 4, \nu_{k25} = +1.0;$
- Resource $r = 3$ ($10te$ glass-lined reactors)
  $\mu_{k30} = -1, \mu_{k3\theta} = 0, \theta = 1, \ldots, 4, \mu_{k35} = +1; \nu_{k3\theta} = 0, \theta = 0, \ldots, 5;$
- Resource $r = 4$ (steam)
  $\mu_{k4\theta} = 0, \theta = 0, \ldots, 5; \nu_{k40} = -0.25, \nu_{k41} = +0.25, \nu_{k4\theta} = 0, \theta = 2, \ldots, 5,$
- Resource $r = 5$ (cooling water)
  $\mu_{k5\theta} = 0, \theta = 0, \ldots, 5; \nu_{k51} = -2.0, \nu_{k55} = +2.0, \nu_{k5\theta} = 0, \theta \neq 1, 5.$

As usual with discrete time formulations, changes to the resource utilisation can only occur at time interval boundaries. We define an additional set of variables, R_{rt} to denote the amount of *excess resource* r over time interval t, *i.e.* the amount of r which is *not* being used by tasks that are active over this interval.

[3] The term "consume" can also be interpreted as "engage" or "reserve".
[4] The term "produce" can also be interpreted as "disengage" or "release".

A Unified Mathematical Formulation for Process Scheduling

The process scheduling formulation based on the RTN concept as introduced above involves only three different classes of constraints in terms of the three types of variables, namely N_{kt}, ξ_{kt}, and R_{rt}.

Excess Resource Balances

These express the fact that the amount of excess resource of a given type changes from one time interval to the next due to the interactions of this resource type both with active tasks and with the environment. They are of the form:

$$R_{rt} = R_{r,t-1} + \sum_{k} \sum_{\theta=0}^{\tau_k} \left(\mu_{kr\theta} N_{k,t-\theta} + \nu_{kr\theta} \xi_{k,t-\theta} \right) + \Pi_{rt} \qquad \forall\, r,\, t \tag{2}$$

where Π_{rt} is the amount of resource r made available from external sources at time t. A negative value would denote an amount removed from the plant at time t. The quantity R_{r0} corresponds to the amount of resource r that is available initially, and is assumed to be given.

Excess Resource Capacity Constraints

The amount of excess resource at any given time cannot be negative. There may also be an upper bound on the amount of resource that can be stored over any interval (*e.g.* for material resources, hot water *etc.*). This leads to the bounds constraints:

$$0 \le R_{rt} \le R_{rt}^{\max} \qquad \forall\, r,\, t \tag{3}$$

Operational Constraints

These may be used to impose certain restrictions on the feasible operation of a task, *e.g.* by limiting the relative magnitudes of the discrete and continuous extents. One common constraint of this type is that of minimum and maximum batch size with respect to processing equipment capacity, which can simply be written as:

$$V_{kr}^{\min} N_{kt} \le \xi_{kt} \le V_{kr}^{\max} N_{kt} \qquad \forall\, k,\, t,\, r \in PE_k \tag{4}$$

where PE_k is the set of processing equipment resources used by task k, with $V_{kr}^{\min}$ and $V_{kr}^{\max}$ being the corresponding minimum and maximum useful equipment capacities.

Finally, the *objective function* will typically be a linear combination of the variables N_{kt}, ξ_{kt}, and R_{kt}.

Comparison with STN-based Formulations

The simple formulation presented above covers entirely all features of the short-term scheduling problem as studied by Kondili *et al.* (1993). In particular, the resource balance constraints (2) replace the constraints expressing processing unit allocation, material balance and utility utilisation in the STN-based formulation presented by Kondili *et al.*

It is also interesting to note that the two formulations are entirely equivalent mathematically under certain conditions. The main binary variables in the STN formulation are denoted by W_{ijt} and take a value of 1 if the STN task i starts being performed in processing unit j at time t. Now consider the case in which each equipment item is treated as a distinct resource type in the RTN, with each RTN task k corresponding to a task/unit pair (i, j) in the STN. Then the integer variables N_{kt} can only take binary values since no more than one instance of any task k can start at any given time. Moreover, the number of such variables in the RTN formulation is exactly the same as the number of W_{ijt} variables in the STN one.

Furthermore, it can be shown that the equipment allocation constraints in the STN formulation, expressed as (Shah *et al.*, 1993a):

$$\sum_{i} \sum_{\theta=0}^{\tau_i - 1} W_{ij,t-\theta} \leq 1 \quad \forall\, j, t \tag{5}$$

where τ_i is the duration of STN task i, can, in fact, be transformed through a sequence of simple manipulations to the RTN constraints (2). Thus, by introducing a slack variable W_{0jt} in (5), we convert it to an equality:

$$\sum_{i} \sum_{\theta=0}^{\tau_i - 1} W_{ij,t-\theta} + W_{0jt} = 1 \quad \forall\, j, t \tag{6}$$

Now, subtracting (6) written at time $(t-1)$ from its counterpart at time t, we obtain:

$$W_{0jt} = W_{0j,t-1} + \sum_{i \in I_j} W_{ij,t-\tau_i} - \sum_{i \in I_j} W_{ijt} \quad \forall\, j, t \tag{7}$$

which is exactly in the form of the RTN resource balance constraint (2), with the slack W_{0jt} acting as the excess resource variable.

Similar manipulations may be used to transform the STN utility utilisation constraints to the balance form (2), while the STN mass balance constraints are already in this form. Hence, overall the two formulations are equivalent mathematically in the sense that they have the same number of binary variables, the same linear programming relaxations and the same integrality gap.

In the case of production scheduling, the main computational advantage of the RTN formulation over its STN counterpart arises in problems involving many identical processing equipment items. Here, the RTN formulation introduces a single integer variable instead of the multiple binary variables used by the STN formulation, and this results both in smaller linear programming relaxations, and in reduced integer degeneracy in the solution of the MILP. A second, albeit smaller, computational advantage arises in cases involving long processing tasks, in which the balance constraints (2) are often much sparser than their STN counterparts (*cf.* constraint (5)).

Further features of the production scheduling problem, as well as other process scheduling problems can be accommodated within the RTN formulation simply by identifying the relevant resources and tasks. For instance, sequence-dependent cleaning requirements can easily be imposed by treating processing equipment in different states of cleanliness as different resources, with the various processing and cleaning tasks effecting appropriate resource transformations. The scheduling of pipeless plants can also be described entirely in terms of the three constraints (2)-(4) simply by identifying the processing stations, the transferable vessels and the AGVs at the various locations with appropriate resource types, and introducing the tasks describing the necessary processing and transfer operations.

Concluding Remarks

This paper has argued that the establishment of unified frameworks for treating a wide variety of process scheduling problems is both desirable and feasible. Two such frameworks, respectively based on the State-Task Network and Resource-Task Network process representations, have been presented.

The STN framework has evolved over several years and has been extended to produce a number of formulations covering many process scheduling problems. The size of the mixed integer mathematical programming problems resulting from the application of this approach means that not all practical problems that one would like to address are solvable in this way at present. On the other hand, it should also be recognised that a large number of industrial scheduling problems — including many whose complexity makes even their description impossible with earlier techniques — are already well within the scope of the approach. Continuing advances in computer hardware and in linear and mixed integer linear programming algorithms will further enhance its applicability.

The RTN framework is a much more recent development. For the case of simple production scheduling problems, it leads to formulations that can be shown to be mathematically equivalent to those obtained from the STN framework. Its main advantage over the latter perhaps lies in its conceptual simplicity and its direct applicability to a large number of complex process scheduling problems beyond simple production scheduling applications. In essence, all process scheduling problems viewed within the RTN framework are equivalent, and can be treated by a common mathematical formulation.

Another major consequence of the simplicity of the RTN framework is the relative ease with which it can be extended to allow the use of more complex descriptions of process operations. In the simplest case, more detailed task models can be accommodated within the context of the mathematical formulation presented in this paper by allowing the extent variables N_{kt} and ξ_{kt} to be *vectors* of decision variables instead of scalars. This would permit, for instance, the modelling of transportation of mixed loads by lorries in multisite plant scheduling: in this case, each element of the continuous extent vector ξ_{kt} would correspond to a different material being transported within the same lorry. The operational constraint (4) would have to be modified to ensure that the total lorry capacity is not exceeded.

More fundamentally, the unified treatment of all resources by the RTN framework could facilitate the development of alternative mathematical formulations, such as for instance, formulations based on continuous representations of time, or incorporating descriptions of uncertainty in the availability of resources.

Acknowledgement

The author is indebted to Dr Nilay Shah and Dr Matthew Realff for many useful discussions relating to some of the material presented in this paper.

References

Birewar, D.B., and I.E. Grossmann (1990). Simultaneous Synthesis, Sizing, and Scheduling of Multiproduct Batch Plants. *Ind. Eng. Chem. Res.*, 29, pp. 2242-2251.

Corominas, J., A. Espuña, and L. Puigjaner (1993). A New Look at Energy Integration in Multiproduct Batch Processes. *Computers chem. Engng.*, 17S, pp. S15-S20.

Crooks, C.A., K. Kuriyan and S. Macchietto (1992). Integration of Batch Plant Design, Automation and Operation Software Tools. *Comput. chem. Engng.*, 16S, pp. S289-S296.

Crooks, C.A., N. Shah, C.C. Pantelides, and S. Macchietto (1993). Detailed Scheduling of General Batch and Semi-Continuous Process Operations. *Ind. Eng. Chem. Res.*, submitted for publication.

Dedopoulos, I.T., and N. Shah (1993). *Optimal Short-term Scheduling of Maintenance and Production for Multipurpose Plants.* Technical Report, Centre for Process Systems Engineering, Imperial College, London, United Kingdom.

Kondili, E., C.C. Pantelides, and R.W.H. Sargent (1988). A General Algorithm for Scheduling Batch Operations. *Proc. Third Intl. Symp. Proc. Syst. Engng.*, Sydney, pp. 62-75.

Kondili, E., C.C. Pantelides, and R.W.H. Sargent (1993). A General Algorithm for Short-Term Scheduling of Batch Operations—I. MILP Formulation. *Computers chem. Engng.*, 17, pp. 211-227.

Mauderli, A., and D.W.T. Rippin (1979). Production Planning and Scheduling for Multi-purpose Batch Chemical Plants. *Computers chem. Engng.*, 3, pp. 199-206.

Niwa, T. (1991). Transferable Vessel-type Multi-purpose Batch Processes. *Proc. Fourth Intl. Symp. Proc. Syst. Engng.*, Montebello, Canada.

Pantelides, C.C., M.J. Realff, and N. Shah (1992). Short-term Scheduling of Pipeless Batch Plants. *AIChE National Meeting*, Miami Beach, Miami, November.

Papageorgiou, L.G., and Pantelides, C.C. (1993). A Hierarchical Approach for Campaign Planning of Multipurpose Batch Plants. *Computers chem. Engng.*, 17S, pp. S27-S32.

Papageorgiou, L.G., N. Shah, and C.C. Pantelides (1992). A Software System for Optimal Planning and Scheduling of General Batch Operations. In *Advances in Process Control III*, The Institution of

Chemical Engineers, Rugby, England, pp. 161-170.
Papageorgiou, L.G., N. Shah, and C.C. Pantelides (1993a). Energy Integration in Multipurpose Plant Operation. Paper presented at CHISA'93 conference, Prague, Czech Republic, August/September.
Papageorgiou, L.G., N. Shah, and C.C. Pantelides (1993b). *gBSS — Optimisation Software for Multipurpose Plants — User Manual and Language Reference*, Technical Report, Centre for Process Systems Engineering, Imperial College, London, United Kingdom.
Rapacoulias, C., N. Shah, and C.C. Pantelides (1991). Optimal Scheduling of Order-Driven Batch Chemical Plants. In *Computer-Oriented Process Engineering*, (L. Puigjaner and A. Espuña, Eds.), pp. 145-160, Elsevier, Amsterdam.
Reklaitis, G.V. (1992). Overview of Scheduling and Planning of Batch Process Operations. *Proc. NATO Advanced Study Institute on Batch Processing Systems Engineering*, Antalya, Turkey.
Rippin, D.W.T. (1993). Batch Process Systems Engineering: a Retrospective and Prospective Review. *Computers chem. Engng.*, S17, pp. S1-S13.
Shah, N. (1992). *Efficient Scheduling, Planning and Design of Multipurpose Batch Plants.* PhD Thesis, University of London.
Shah, N., and C.C. Pantelides (1991). Scheduling of Lubricants Blending Operations — a Case Study in Optimal Scheduling. *AIChE Annual Meeting*, Los Angeles, California, November.
Shah, N., C.C. Pantelides, and R.W.H. Sargent (1993a). A General Algorithm for Short-Term Scheduling of Batch Operations—II. Computational Issues. *Computers chem. Engng.*, 17, pp. 229-244.
Shah, N., C.C. Pantelides, and R.W.H. Sargent (1993b). Optimal Periodic Scheduling of Multipurpose Batch Plants. *Annals of Operations Research*, 42, pp. 193-228.
Shimatami, T. and O. Okuda (1992). 'Pipeless' Batch Chemical Plants Offer a New Approach. *Chem. Engng.*, October.
Takahashi, K. and H. Fujii (1990). New Concept for Batchwise Specialty Chemical Production Plant. *Intl. Chem. Engng.*, September/October.
Vaselenak, J.A., I.E. Grossmann, and A.W. Westerberg (1986). Heat Integration in Batch Processing. *Ind. Eng. Chem. Proc. Des. Dev.*, 25, pp. 357-366.
Wellons, M.C., and G.V. Reklaitis (1989a). Optimal Schedule Generation for a Single-Product Production Line, Part I: Problem Formulation., *Computers chem. Engng.*, 13, pp. 201-212.
Wellons, M.C., and G.V. Reklaitis (1989b). Optimal Schedule Generation for a Single-Product Production Line, Part II: Identification of Dominant Unique Path Sequences. *Computers chem. Engng.*, 13, pp. 213-227.
Wellons, M.C., and G.V. Reklaitis (1991a). Scheduling of Multipurpose Batch Chemical Plants. 1: Formation of Single Product Campaigns. *Ind. Eng. Chem. Res.*, 30, pp. 671-688.
Wellons, M.C., and G.V. Reklaitis (1991b). Scheduling of Multipurpose Batch Chemical Plants. 2: Multiple Product Campaign Formation and Production Planning. *Ind. Eng. Chem. Res.*, 30, pp. 688-705.
Zanetti, R. (1992). The 'Pipeless' Batch Plant. *Chem. Engng.*, June.