Journal of Intelligent Manufacturing (1997) **8**, 425–433

On the receding horizon hierarchical optimal control of manufacturing systems

F. LOBO PEREIRA and J. BORGES DE SOUSA

Instituto de Sistemas e Robótica and DEEC–Faculdade de Engenharia da Universidade do Porto, R. dos Bragas, 4099 Porto Codex, Portugal

Received May 1997 and accepted June 1997

This paper concerns the development of a hierarchical framework for the integrated planning and scheduling of a class of manufacturing systems. In this framework, dynamic optimization plays an important role in order to define control strategies that, by taking into account the dynamic nature of these systems, minimize customized cost functionals subject to state and control constraints. The proposed architecture is composed of a set of hierarchical levels where a two-way information flow, assuming the form of a state feedback control, is obtained through a receding horizon control scheme. The averaging effect of the receding horizon control scheme enables this deterministic approach to handle random and unexpected events at all levels of the hierarchy. At a given level, production targets to the subsystems immediately below are defined by solving appropriate optimal control problems. Efficient iterative algorithms based on optimality conditions are used to yield control strategies in the form of production rates for the various subsystems. At the lower level, this control strategy is further refined in such a way that all sequences of operations are fully specified. The minimum cost sensitivity information provided in the optimal control formulation supports a mechanism, based on the notion of a critical machine, which plays an important role in the exploitation of the available flexibility. Finally, an important point to note is that our approach is particularly suited to further integration of the production system into a larger supply chain management framework, which is well supported by recent developments in hybrid systems theory.

Keywords: Hierarchical optimal control, manufacturing systems, receding horizon

1. Introduction

The wide range of extremely difficult and unsatisfactorily answered issues arising in the design and development of control and planning architectures for CIM systems, makes it a challenging task subject of active research in spite of the strong effort that have been undertaken in the past decades by a large R&D community, (Gershwin *et al.*, 1984; Friedman *et al.*, 1992). This effort has been strongly motivated by the important role that optimal planning and scheduling of manufacturing systems play in the competitive global environment that most industries are facing today.

A significant body of R&D work searching for optimization of manufacturing policies deals with frameworks whereby the typical problem (French, 1982; Hu and Caramanis, 1992) of scheduling a fixed number of jobs with known processing requirements on a given set of machines so that some typical performance measure is minimized, is usually formulated as a combinatorial optimization problem. Although an efficient approach based on a Lagrangian relaxation technique is presented in Luh and Hoitomt (1993), typically, solution methods suffer from the 'curse of dimensionality' and require the consideration of heuristic algorithms which can yield good solutions from a mathematical point of view. The major drawback of these approaches lies in the fact that they do not involve an adequate description that takes into account the true dynamic nature of the addressed manufacturing systems. These may be viewed as discrete event dynamical systems whose state evolution, governed by several decision layers, occurs in different time scales and is affected by undesirable perturbations (Gershwin, 1987, 1989; Sousa, 1991; Sousa and Pereira, 1992; Chauvet *et al.*, 1997). For this reason, a paradigm dynamical optimization enabling the integration of planning and scheduling activities constitutes a promising framework. Furthermore, the integration of the decision-making structure of the production system within a larger framework encompassing the relevant supply chain is an important consideration strengthening the links with

0956-5515

the strategic management of the enterprise (Leach *et al.*, 1996; Pfohl, 1996).

This work was motivated by the analysis of two different classes of production systems having in mind the respective integrated planning and scheduling. One is a high-volume semiconductor assembly factory where the propagation of work-in-progress inventories through the system causes undesirable instabilities (Sousa, n.d.). The other is a textile company composed of spinning, dyeing and weaving units whose coordinated control, involving an accurate forecast of delivery times, represents a major difficulty (Gonçalves *et al.*, 1997). This analysis revealed that the set of control problems arising in both classes of production systems was essentially the same and suggested a general framework which has been gradually maturing.

The solution approach to these problems involves an issue of primary interest consisting of the definition of the required control space by specifying a structure of buffers preceding some groups of machines which are effective in imposing the required behaviour. We have proposed a hierarchical control structure (Sousa and Pereira, 1992, 1994) that, for a given time horizon, and given demand and supply constraints profile, permits us to find a robust and least suboptimal admissible schedule for the manufacturing system.

In order to reduce the complexity and problem size underlying the hierarchic decision-making structures of these production systems, which involve an extremely large number of machines, product-types and operations, a method combining an adapted aggregation of parts and machines with the use of continuous measures of the occurrence of discrete events was adopted. The higher the level of the hierarchy, the longer is the considered time horizon and the greater is the degree of aggregation of parts and machines performed in such a way as to find the best compromise between complexity and model accuracy. In this way, tractable mathematical descriptions of the dynamic behaviour are obtained and taken into account in the formulation of customized optimal control problems. It is also important to note that the adopted adaptive receding horizon scheme contributes to minimize the adverse effects in the control policy due to random events, model inaccuracy (e.g., set-ups) and the mismatch between models at the different levels of the hierarchy.

Finally, another point to note concerns the fact that an integrated decision support system was designed for a spinning unit of a textile company (Gonçalves *et al.*, 1997) based on the proposed control framework. In order to satisfy the proposed objectives and constraints, a systems engineering process was carried out to reveal the functional architecture, which was mapped on a three-level conceptual architecture.

The remainder of this paper is organized as follows. After commenting on related work in Section 2, we present a detailed description of our approach in Section 3. Here, special attention is paid to the modelling approach leading to the optimal control problem formulation. In Section 4, the solution method is explained in detail. Then a receding horizon algorithm incorporating the higher and lower level solution methods is presented. In Section 5, a brief reference is made to the methodology adopted in order to design the decision support system for the spinning unit of a textile company using the proposed framework as a conceptual basis. Finally, in the last section, we state some concluding remarks and point out directions for future research.

2. Related work

Several researchers have used various control theoretical approaches to address the production planning and scheduling problem.

Concerned with computational tractability and the definition of control policies based on realistic models, Kimenia and Gershwin (1983) developed a pioneering approach to the production control of manufacturing systems with failure-prone machines. A multilevel hierarchical control algorithm, involving stochastic optimal control at the first level, was proposed. The control variables of this problem are the production rates for the whole system. In this way, the behaviour of the system is captured in an aggregated form, since the huge number of occurrences of discrete events is replaced by a small number of real continuous-time functions, representing the rates at which they occur. The computation of those rates, in the form of a feedback control law, by considering a quadratic approximation for the corresponding cost-to-go function and a tracking policy, the staircase policy, was proposed in order to transform the continuous production rates into the discrete loading times of jobs into the system. The notion of 'hedging point' was also introduced to designate the optimal buffer level with which to hedge against future failures. This framework was further extended in Gershwin (1987, 1989) in order to cope with the multiple time scales related to the occurrence of desirable (production activities) and undesirable (set-ups and machine failures) events. The system is seen as a set of hierarchical levels (where each level corresponds to events taking place at a given frequency) and the approach consists of finding an at least suboptimal frequency for each controllable activity subject to appropriate constraints imposed by the higher-level activities.

Eleftheriu and Desrochers (1987) have extended the concept of the hedging point developed by Kimenia and Gershwin (1983), in order to find a hierarchical controller that minimizes product inventory surplus and backlog costs while keeping production rates as close as possible to demand. Perkins and Kumar (1989) have defined scheduling policies which ensure the stability of inventory in systems where machines are not entirely flexible. Akella *et al.*,

(1990) have developed a hierarchical control scheme for a failure-prone parallel multicell flexible assembly line. Then, based on a tractable model obtained by assuming a number of simplifying assumptions, a stochastic optimal control problem is formulated and a solution method is proposed. A quadratic approximation to the cost function is used in order to find a computational solution to the problem.

The work of Sharifnia (1992), Connors *et al.*, (1992) and Egilmez and Sharifnia (1994) is particularly interesting since it relates directly to our approach (Sousa, n.d.; Sousa and Pereira, 1992, 1994) by using the continuous flow production models. Sharifnia obtained a linear formulation for the continuous flow approximation of the system and proposed to re-solve the problem whenever significant random disturbances or deviations from the optimal continuous targets were observed. Connors *et al.* have focused on myopic control policies which require some important simplifications to the original problem formulation and, consequently, do not necessarily yield an optimal solution. In the framework of continuous flow approximations it is particularly difficult to model set-up operations. Using two different approaches, Srivatsan and Gershwin (1991), Hu and Caramanis (1992) and Bai and Elhafsi (1994) obtained an interesting description of the optimal set-up behaviour of certain manufacturing systems. In the last reference, the characterization of the optimal solution to the problem of scheduling a single deterministic machine with a two-part-type set-up revealed the complexity underlying this very simple problem.

Although in a static optimization framework, Luh and Hoitmont (1993) have used an interesting interpretation of Lagrange multipliers as a vector of prices in order to achieve the coordination of solutions to a set of subproblems obtained by Lagrangian relaxation. This plays a role analogous to our adjoint variable associated with the solution to the optimal control problem. Recognizing the N-P hard complexity of the problem of optimizing manufacturing schedules, Chu (1996) has proposed a unified framework of a different character. Knowledge acquired by a neural network during the running of the production system is used in order to select the best scheduling algorithms for a given context.

In Chauvet *et al.* (1997), a unified approach to the hierarchical management and control of production systems developed in the environment of algebra and factory dynamics was presented. This approach leads to a new hierarchical structure and a new aggregation concept where the modularity of dynamic and decision models within all levels of the hierarchy constitute the unifying link.

3. Proposed approach

Starting with the basic description of the manufacturing system, this section addresses the hierarchic control structure composed of several interconnected decision-making problems at various levels. Particular attention will be drawn to the amenability of the proposed structure w.r.t. the integration in a wider context of supply chain management.

3.1. *Manufacturing system description*

In order to describe the model of the manufacturing system the following notation is required:

N – time units of the planning horizon;
m – machine-type index, $m = 1, \dots, M$;
n_m – number of type m machines operating in parallel;
i – product-type index;
τ_m – vector of processing times at a machine of type m with components $\tau_{m,i}$;
U_m – production rates constraint set of type m group of machines;
V_P – product input rate constraint set;
Ω $= U_1 \times \cdots \times U_M \times V_P$ – the global control constraint set.

At time t:

$u_m(t)$ – vector of production rates for the group of machines of type m with components $u_{m,i}(t)$;
$v_p(t)$ – vector of the input rate of products of type p to the system;
$u(t)$ $= \mathrm{col}(u_1(t), \dots, u_M(t), v_1(t), \dots, v_P(t))$;
$b_m(t)$ – vector of buffer contents for the group of machines of type m with components $b_{m,i}(t)$;
$b(t)$ $= \mathrm{col}(b_1(t), \dots, b_M(t))$;
$d(t)$ $= \mathrm{col}(d_1(t), \dots, d_P(t))$ – vector of the demand profile for the system.

Here, col(...) is a column vector with entries ordered as presented between parentheses, $a \cdot b$ is the inner product of vectors a and b, ϕ is the null vector, and $\mathrm{A} \times \mathrm{B}$ denotes the Cartesian product of sets A and B.

The class of manufacturing systems we address in this paper is generally described by the following features:

(1) There are P and M product and machine types respectively;

(2) The entity machine-type m is characterized by n_m machines operating in parallel (the same regime). If two identical machines are not coupled then they should be included in different machine types;

(3) Discrete part and continuous processing is allowed within the same system's model;

(4) Assembly and disassembly of discrete products are considered;

(5) To each product type there corresponds a set of alternative routings or flows. Each routing specifies a sequence of operations at pre-defined machines. Reentrant flows may be considered;

(6) Operation in the system are specified by product and machine types with known processing times;

(7) Multiple visits to a machine, by the same product, are permitted;

(8) Some machines are entirely flexible, while others may require a deterministic set-up time during transition between two set-up states;

(9) Demand profiles, $d_p(t)$, are known *a priori*;

(10) It is permitted to build inventories before some machines. The inventory associated with a machine is denoted as the buffer of the machine.

These hypotheses imply that the flow control modelling approach provides a good approximation to the behaviour of the manufacturing system by using continuous production rates to describe the occurrence of discrete events. It is clear that the quality of the approximation is directly proportional to the inverse of the highest processing time in the system. In particular, this approach is also amenable to incorporate continuous processing.

The existence of a buffer before a certain machine extends the set of feasible solutions and permits greater controllability, since products can be held for some time. It is justified in order to:

(1) Adapt production rates from different machines;

(2) Ensure the full utilization of bottleneck machines;

(3) Build inventories to support types of generalized hedging point strategies to deal with future failures or set-ups in feeding machines;

(4) Uncouple the behaviour of machines;

(5) Provide the means for the system to respond to demand which may instantaneously exceed the available capacity.

3.2. *The optimal control problem*

In this section we describe an instance of the various constraints which have to be satisfied by the control strategy while minimizing the cost functional. The consideration of set-up times would make the model much more complex and it would obscure the main idea of this paper. Therefore, we assume they are negligible compared to the production times and consider the corresponding machines entirely flexible.

Capacity constraints depend not only on the intrinsic properties of isolated machines but also on the system's operating point and therefore they will appear naturally in the system's dynamic equations.

It is relatively straightforward to show that, at time t, the vector of production rates $u_m(t)$ corresponding to a group of n_m flexible machines of type m satisfies the following inequalities:

$$\tau_m \cdot u_m(t) \le n_m \quad \forall\, m \tag{1}$$

$$u_m(t) \ge \phi \tag{2}$$

where this inequality should be understood componentwise. This set of feasible production rates assumes the form of a convex polyhedron (Sousa, n.d.). Machines with similar operating properties and not separated by buffers may be grouped so that their joint behaviour is completely described by an unique equivalent machine whose capacity constraint is also a convex polyhedron that corresponds to the intersection of k convex polyhedrons (Sousa, n.d.).

In this problem the vectors of buffer content $b(t)$ and production rates $u(t)$ play the role of state and control variables respectively. A particular instance of a discrete-time version of the optimal control problem requiring the demand to be necessarily satisfied within the due date may be stated as follows:

$$\text{Minimize} \sum_{t=1}^{N}\left\{ {c_1}^T b(t) + {d_1}^T\left[\sum_{k=0}^{t} O \cdot u(k) - d(k)\right]\right\} \tag{3}$$

subject to:

$b(0) = b_0$
$b(t+1) - b(t) = Au(t)$ (dynamic equation)
$b(t) \ge \phi$ (nonnegative buffer content)
$u(t) \in \Omega_t$ (capacity constraints)
$\sum_{k=0}^{t}[O \cdot u(k) - d(k)] \ge 0$ (demand satisfaction)

where the last four relationships should hold for all values of t and are understood componentwise. A is a matrix specifying the flow of parts, and O indicates outbound products which may include semifinished parts. Note that, for the inbound components of u, Ω_t may represent the supply constraints whose profile may vary with time in a either deterministic, random or 'controlled' way. The third constraint not only prevents the buffer content from becoming negative but also represents the capacity dependence on the operating point. This situation arises when the buffer between adjacent machines is empty and it is not possible to input to the system the relevant semifinished part. In this situation the production rates of those machines are equal and limited above by the capacity of the slowest machine.

The first component of the cost function (Equation 3) penalizes the buffer content in order to decrease the cycle time for all products. The second component contributes to the reduction of the excess of output w.r.t. demand. The cost functional could also incorporate other typical objectives of manufacturing management organizations, such as:

(1) Penalization of sudden changes of the system configuration or backlogs;

(2) Preservation of the flexibility of the current decision, in the sense that the decision made at the present time keeps options open in the future. In Lasserre and Roubelat (1985), measures of flexibility are presented.

Obviously, a multipart machine set-up could be included in the model of the system dynamics. In this case some continuous approximation to the set-up operation is required in order to ensure the regularity of the state dependence of the dynamics required by the applicability of the optimality conditions. We remark that, when the set-up time is negligible w.r.t. to each single part operation time, the added complexity (e.g., Bai and Elhafsi, 1994), together with the additional analysis to ensure the validity of the limiting procedure of the approximation scheme may bring little value in the face of naturally occurring random events as well as model mismatch inherent to the aggregation procedures. However, when the set-up operations are significant, then a hybrid automata model (Alur *et al.*, 1993) should be adopted where different dynamics are considered for each one of the discrete states. We will not dwell on this point, but we just mention that the different states may correspond to the various types of set-ups and of multiple part configurations of the machine and that the discrete state transitions are part of the control solution which, in this framework, are obviously subject to optimization. It should also be noted that a hybrid model is needed whenever a major reconfiguration is in order, either imposed by a major disruption, or as a result of a strategic decision in order to improve the overall system's performance.

Although this problem is not in a form to which our solution method based on the maximum principle (Halkin, 1966) might be applied directly, it is straightforward to conclude that the following change of variables:

$$\begin{aligned}
e(t) &= \sum_{k=0}^{t} O \cdot u(k), \quad \bar{d}(t) = \sum_{k=0}^{t} d(k), \quad t = 0, \ldots, N \\
e(0) &= Ou(0), \quad \bar{d}(0) = d(0), \quad \bar{u}(t) = \mathrm{diag}(u(t), u(t+1)) \\
\bar{b}(t) &= \mathrm{col}(b(t), e(t)); \quad \bar{A} = \mathrm{diag}(A, O) \\
\bar{G} &= \mathrm{diag}(I, I); \quad d^*(t) = \mathrm{col}(\phi, \bar{d}(t)) \\
\bar{c}(t) &= \mathrm{col}(c, d_1)
\end{aligned}$$

being I and ψ, respectively, the identity and the null matrices, ϕ, the null vector and col(A, B) and diag(A, B), respectively, column block and diagonal block matrices with components A and B, the original problem is, for feasible demand profiles, converted into the following standard problem:

Minimize $\sum_{t=1}^{N}[\bar{c}^T\bar{b}(t) + {d_1}^T\bar{d}(t)]$ subject to

$$\bar{G} \cdot \bar{b}(0) = \mathrm{col}(b_0, Ou(0)) \tag{4}$$

$$\bar{b}(t+1) - \bar{b}(t) = \bar{A}u(t), \quad t = 0, \ldots, N-1 \tag{5}$$

$$\bar{G} \cdot \bar{b}(t) \geq d^*(t) \quad t = 0, \ldots, N \tag{6}$$

$$\bar{u}(t) \in \Omega_t \times \Omega_{t+1}$$

In the next section is presented a method to compute the solution of this problem which establishes the production rates for all the machines of the system. These are the inputs to the lower level of the hierarchy in the form of production targets that will enable the generation of the schedule for all the machines of the system.

Before closing this section, three remarks should be made w.r.t. to the problem formulation.

The first one concerns the fact that we considered vectors c and d_1 as fixed per part × time-unit costs incurred in maintaining inventory. However, a different meaning of per-unit value could be attached to these coefficients. These parameters could be part of the decision making and defined at the upper level if a wider framework enabling the integration of the production system in a supply chain management system is considered.

The second one concerns the fact that this framework encompasses other problem formulations such as (those resulting from): withdrawing the constraint of Equation 6 and penalizing the tardiness in the objective function; introducing additional non-controllable controls in the dynamics to take into account random events, set-up effects on capacity, model mismatch introduced by the aggregation procedure, etc. In this last formulation, the solution method would seek to optimize the controllable inputs for the worst case of the non-controllable ones.

Finally, the third remark concerns the dynamics which could be given by another type of state transition scheme, as far as are retained semi-group properties with at least a Lipschitz continuity dependence on the state variable.

4. Solution method

In order to keep the exposition clear, we will consider a two-level hierarchy which, for simpler systems, might be reasonable and which captures all the essential ingredients of the presented approach. The approach to solve the whole problem consists of embedding the two level hierarchy of solution stages in a receding horizon control framework. The higher-level stage generates production rate guidelines for real-time operation, dealt with by the lower-level stage. In the following we will describe the three components of the solution method.

4.1. *Higher-level stage*

An algorithm developed in Pereira *et al.* (1991) has been used in order to find the set of optimal production rates which are the output of the first stage. This is an iterative procedure using feasible descent directions to search for a control function satisfying necessary conditions of optimality (Clarke, 1983).

Note that this algorithm uses the maximum principle for optimal control problems in the absence of state constraints such as Equation 6. However, an exact penalization result in Clarke (1983) implies that the given problem has the same solution as another one where state constraints are not present and their violation is accounted for by penalizing the cost functional as follows:

$$\sum_{t=1}^{N}\left\{\bar{c}^T\bar{b}(t) + d_1{}^T\bar{d}(t) + K\,e^T\max[0, d^*(t) - \bar{G}\cdot\bar{b}(t)]\right\} \quad (7)$$

where $e = \text{col}(1,1,...,1)$, the max operation is understood componentwise and $K > 0$ is a sufficiently large constant.

The maximum principle applied to this problem states that if $(\bar{b}(t), \bar{u}(t))$ is an optimal control process then the control policy $u(t)$ maximizes on Ω , for each t, the map:

$$u \rightarrow H(t, \bar{b}(t), p(t), u) \quad (8)$$

The pseudo-Hamiltonian $H(t, \bar{b}(t), p(t), u)$ is defined by:

$$\begin{aligned} &p^T(t+1)[\bar{b}(t) + \bar{A}(t)u(t)] - \{\bar{c}(t)^T\bar{b}(t) \\ &\quad + d_1{}^T\bar{d}(t) + Ke^T\max[0, d^*(t) - \bar{G}\cdot\bar{b}(t)]\} \end{aligned} \quad (9)$$

where the adjoint variable $p(t)$ satisfies:

$$p(N) = 0 \text{ and } p^T(t) = p^T(t+1) - \bar{c}(t)^T - k^T(t) \quad (10)$$

where $k^T(t) = K(\alpha_1, \alpha_2, \ldots, \alpha_z)$, z is the state space dimension, and α_i is given by 1 if $d_i^*(t) > \bar{G}\cdot\bar{b}(t)_i, -1$ if $d_i^*(t) > \bar{G}\cdot\bar{b}(t)_i$, and any point in $[-1, 1]$ otherwise.

Once the initial buffer contents $b(0) = b_0 \in R^Z$ and the weighting factors $\bar{c} \in R^z$ are known, the algorithm described below will produce, after a number of iterations, a control policy yielding a local minimum. Note that in this case, these conditions are also sufficient. The basic version described in Pereira *et al.* (1991) is:

Step 0. Initialization

Set $i := 0$ (iteration counter).

Choose a feasible control policy, u_i.

Step 1. Computation

Compute:

$\bar{b}_i$, by using Equations 3, 5 and 4; p_i, by using Equation 4;

$\nabla_u H_i(u_i)$, the gradient of the pseudo-Hamiltonian (defined by Equation 9) with respect to the control;

$C_{fd}(u_i)$, the cone of feasible control directions at u_i.

Step 2. Test of optimality

If $\nabla_u H_i(u_i) \in C^0_{fd}(u_i)$, the negative polar of $C_{fd}(u_i)$, then the optimal policy is found; stop.

Step 3. Update of control

Let $u_{i+1} = P_\Omega\,(u_i - s_i\nabla H_i(u_i))$, where $P_\Omega(z)$, the projection of z on Ω, is the point in Ω closest to z and s_i is the solution to the one-dimensional subproblem

$$\text{Minimize } [J(P_\Omega(u_i\cdot s\nabla H_i(u_i))) : s > 0]$$

Here $J(u)$ is the integral function in Equation 7 expressed in its total dependence on the control.

Step 4

Set $i = i + 1$ and go to step 1.

By using relatively standard arguments (Pereira *et al.*, 1991) it is concluded, that, under fairly standard hypothesis on the data, this algorithm produces a strictly decreasing sequence of controls converging to a minimizer.

Note that the constant K mentioned above is not known *a priori*. One way to overcome this difficulty consists of starting with a relatively small value and, as iterations proceed, increasing it in a multiplicative fashion.

Since this algorithm is based on the maximum principle which yields conditions of a local variational character, convergence is only ensured when the considered *a priori* estimate is in a relatively small neighbourhood of the solution. However, this condition is fulfilled in our framework, since this optimal control problem is to be solved every time there is a relatively small shift of the time horizon in the implementation of the receding horizon scheme and the initial optimal estimate for the above algorithm is taken as the solution to the previous problem.

In Pereira and Sousa (1993) another algorithm based on the Pontryagin-type necessary conditions of optimality for differential inclusion problems was presented. By allowing the search for optimality without using gradients w.r.t. the control variable, it is better suited to deal with nonsmooth problems.

4.2. *Lower-level stage*

The decision problem at this stage consists of defining for each subinterval of the time horizon fully detailed discrete parts release, routing and machine scheduling decisions so that the aggregated production targets given in the form of production rates by the higher level problem are met and, simultaneously, the best use of the available resources is made. In order to accomplish this function, a wide variety of techniques ranging from optimization procedures (Luh and Hoitomt, 1993) to fixed policies and reactive logical rules (French, 1982), as well as combinations of these, is available, and the nature of the specific subsystem being considered dictates its selection.

However, since the optimal control formulation provides minimum cost sensitivity information via the adjoint variable, this should be used in the algorithm described below so that the adverse effects of random events, aggregation model mismatch and model uncertainty, are minimized.

Step 1

Define a list of the machines ordered by their criticality. As we mentioned before, the more critical the machine the more it contributes to the cost increase when subject to admissible perturbations at the optimal nominal trajectory. In Vinter and Wolenski (1994), it is shown that the adjoint variable, defined in the statement of the maximum principle, provides this valuable piece of information.

Step 2

Concentrate first on the scheduling of all the most critical machines in such a way that the sequence of operations is globally compatible, i.e., satisfying flow and machine constraints. If realistic dynamic constraints are taken into account at the higher-level optimal control problem and the demand profile is feasible, then at least one such sequence exists. Product processing times are used to transform continuous production rates into sequences of discrete events in this process. In critical machines, processing times play the role of variable discretization intervals. When a machine is not working at its full capacity, this variable discretization interval, corresponding to the time gap between the occurrence of consecutive production operations, is determined from the production rates.

Step 3

Allocate and sequence operations in the set of machines with the critical degree immediately below so that the set of all sequenced operations is feasible. Repeat this step until all the operations have been sequenced.

This procedure does not guarantee the uniqueness of the resulting machine schedule and some iterations may be needed to obtain a globally compatible sequence of operations for the whole system. This marginal degree of freedom might be used in order to cope with unexpected events arising in practical situations. For more significant perturbations, the suboptimality margin for the higher level problem should be increased so that a larger degree of flexibility and hence adaptivity is allowed. This type of consideration was discussed with detail in Lasserre and Roubelat (1985).

4.3. *Receding horizon strategy*

We begin by describing this strategy in the form of an algorithm. Denote the optimal control problem with initial conditions (x_0, t_0) and horizon N, as described in Section 3, by $P^N(x_0, t_0)$.

Step 1. Initialization

$i := 0; t_i := t_0, x_i^*(t_i) = x_0 i$ and $T_i = T$.

Step 2. High-level control problem

Solve $P^{T_i}(x_i^*(t_i), t_i)$ for the interval $[t_i, t_i + T_i]$. This yields the optimal process (u_i^*, x_i^*) on $[t_i, t_i + T_i]$.

Step 3. Low-level control problem

Solve this problem on $[t_i, t_f]$, where $t_f \in (t_i, t_i + T_i]$.

Step 4. Preparation of new iteration

Let $i := i+1$, $t_i = t_f$, $x_i^*(t_i) = x_{i-l}^*(t_f)$. Pick a new T_i. Go to step 2.

A key observation on the above algorithm concerns the fact that the maximum principle used to solve the higher level optimal control problem in step 2 yields an open loop control that depends on the initial state. Thus, when computing the solution for the new time subinterval we are in fact closing the loop. The receding horizon feedback control law provided by steps 2 and 3 plays a key role in this algorithm. The notion of receding horizon control has long been known but, only recently, results concerning the robustness and stabilizing properties of this scheme for problems with state-control constraints, general cost function and nonlinear system's dynamics have been proved (Keerthy and Gilbert, 1988; Mayne and Michalska, 1990; Vinter and Michelska, 1991). Conventional receding horizon techniques, such as those described in Keerthy and Gilbert (1988), do not handle some important features of this algorithm:

(1) Problem $P^{T_i}(x_i^*(t_i), t_i)$ is solved only at discrete times, t_i, where the size of $\Delta = t_f - t_i$ depends on the dynamics;
(2) Only approximate solutions to $P^{T_i}(x_i^*(t_i), t_i)$ are needed;
(3) Moderate model discrepancies can be tolerated.

The results in Mayne and Michalska (1990) and in Vinter and Michalska (1991) apply to equivalent versions of our problem where state and control constraints are removed and replaced by adequate penalization terms in the cost functional. Only mild assumptions are required in order to well define the integral cost functional and guarantee the Lipschitz continuity of the value function, which will be employed as a Lyapunov function for the required stability analysis of this scheme.

The duration of time horizon T_i depends on the system behaviour and on the magnitude of random disturbances. To see this, just note that the specification of the demand at the output of the system consists essentially of a pull rule and that the strategy defined by the optimal control problem represents a compromise between utilization of the dynamic capacity of the system and the minimization of product cycle time. Therefore the behaviour of the system at time t depends on the demand at time t plus the maximum cycle time of the products being produced. This way, if the period T_i is a multiple of the maximum cycle time, then the behaviour of the system in the first instant of the current period is not affected by demand after period T_i.

Although the whole framework represents a significant computational burden its structure is particularly suited to take advantage of the computational power offered by distributed and interconnected computer systems of modern manufacturing systems. On the other hand, the receding horizon strategy provides three welcome additional features:

(1) The lower-level problem is solved only for the first subinterval $[t_i, t_f]$ of the horizon $[t_i, t_i + T_i]$;
(2) When mild disturbances are present, the solution of the current finite horizon problem $P^{T_i}(x_i^*(t_i), t_i)$ will not

differ significantly from that of the preceding problem and u^*_{i-1} is a good initial estimate to solve $P^{T_i}(x^*_i(t_i), t_i)$. Furthermore, it is very likely that, for important classes of manufacturing systems, the assumptions validating the 'averaging' approximation of a stochastic problem by a deterministic one (Liptser *et al.*, 1996) are met;

(3) Computational burden resulting from the solution of the optimal control problem is reduced because the T_i is usually shorter than the original planning horizon N.

5. Towards implementation

An issue that is rarely addressed in the literature describing conceptual control architectures concerns the design and implementation issues. Here, we will state the main ideas behind a methodological framework adapted to the specific characteristics of the industrial organization in Portugal that was used to design an integrated decision support system for the spinning unit of a textile company, (Gonçalves *et al.*, 1997). This framework resulted from the application of systems engineering methods that transforms a linguistic expression of objectives and constraints into a system solution. This process includes the following main phases: requirements analysis, functional analysis and system synthesis.

While the first phase yields a set of high-level requirements that will serve as input to the functional analysis, this phase will produce a set of functional requirements expressed in the form of a user-oriented specification to be validated by the system end user. The final phase consists of mapping the obtained functional architecture onto the conceptual architecture and provides, as output, the technical specification for the control architecture implementation.

This process ensures that the information and decision flows required by the various decision-making problems in the architecture are actually available and are meaningful from the point of view of the manufacturing system. Furthermore, the role of humans in the organization is also properly accounted for. This and the fact that the end users have an active role in the analysis and design process have an important impact in the proper use of the capabilities of the designed system.

6. Conclusions

In this paper, we present current developments of an hierarchical approach permitting the definition of suboptimal control strategies for a large class of manufacturing systems. The use of dynamic optimization techniques coupled with a receding horizon scheme provides a framework to tackle planning and scheduling issues encountered in important classes of manufacturing systems.

The fact that it requires a realistic model of the system, the cost functional may be properly customized, and the use of a specific optimization algorithm makes it especially attractive when the quest for optimality is the main concern.

Furthermore, this approach is particularly amenable to integration in a wider context of supply chain management for which models and techniques of hybrid systems theory appears to be the most favourable framework.

The integration of this framework, within a current research project, in a manufacturing environment is in its early stages. Simulation results obtained so far permit us to consider this approach as a promising avenue and much more work remains to be done in the following directions:

(1) Robustness and sensitivity of the receding horizon strategy with respect to larger parameters perturbations;

(2) Full development of an integrated framework supported by the concepts and tools of hybrid systems theory;

(3) Further requirements so that this approach can be used as an efficient tool in a decision support system with particular emphasis in the context of a supply chain management framework;

(4) Implementation issues on a real-time distributed computing system.

Acknowledgements

The authors gratefully acknowledge the financial support of JNICT, Portugal, under Grant PBIC/TRP/1164/92.

References

Akella, R., Krogh, B. and Singh, M. (1990) Efficient computation of coordinating controls in hierarchical structures for failure-prone multi-cell flexible assembly system. *IEEE Transactions on Robotics and Automation*, **6** (6.), 659–672.

Alur, R., Corcubetis, C., Henzinger, T. and Ho, P. (1993) Hybrid automata: an algorithmic approach to the specification and verification of hybrid systems, in *Hybrid Systems*, Grossman, R., Nerode, A., Ravn, A. and Rischel, H. (eds), Lecture Notes in Computer Science no. 736, Springer-Verlag.

Bai, S. and Elhafsi, M. (1994) Optimal feedback control of a manufacturing system with setup changes, in *Proceedings of the Rensselaer Fourth International Conference on CIM*, Troy, NY, pp. 1–6.

Chauvet, F., Canuto, E., Janusz B. and Proth, J.-M. (1997) Hierarchical production control methodology: a unified approach, in *Proceedings of the Second HIMAC Workshop*, Karlsruhe, Germany, 15–16 May, (to appear).

Chu, C. (1996) Production schedule and control: a unified framework, in *Proceedings of the 4th IEEE Media Symposium on New Directions in Control and Automation*, Maleme, Crete, Greece, June pp. 381–385.

Clarke, F. H. (1983) *Optimization and Nonsmooth Analysis*, Wiley, New York.

Connors, D., Feigin, G. and Yao, D. (1992) Scheduling semiconductor lines using a fluid network model, in *Proceeding of the Rensselaer Third International Conference on CIM*, Troy, NY, pp. 174–183.

Egilmez, K. and Sharifnia, A. (1994) Optimal control of a manufacturing system based on a novel continuous-flow model with minimal WIP requirement, in *Proceedings of the Rensselaer Fourth International Conference on CIM*, Troy, NY, pp. 184–191.

Eleftheriu, M. N. and Desrochers, A. A. (1987) A unified approach to production planning and scheduling, in *Proceedings of the 26th IEEE CDC*, pp. 605–611.

French, S. (1982) *Sequencing and Scheduling* Wiley, New York.

Friedman, A., Glim, J. and Lavery, J. (1992) The mathematical and computational sciences in emerging manufacturing technologies and management practices. *SIAM Reports on Issues of Mathematical Sciences*, Philadelphia, USA.

Gershwin, S. B. (1987) A hierarchical framework for manufacturing scheduling: a two machine example, in *Proceedings of the 26th IEEE CDC*, pp. 651–656.

Gershwin, S. B. (1989) Hierarchical flow control: a framework for scheduling and planning discrete events in manufacturing systems. *Proceedings of the IEEE*, **77**(1), 195–209.

Gershwin, S. B., Hildebrand, R., Suri, R. and Mitter, S. (1984) A control theorist's perspective on recent trends in manufacturing systems, in *Proceedings of the 23th IEEE CDC*, pp. 209–225.

Gonçalves, G., Borges de Sousa, J. and L. Pereira, F. (1997) A systems engineering approach to the design of an integrated decision support system for a textile company, in *Proceedings of the International Conference on Robotics and Automation*, Alburquerque, NM, USA, April.

Halkin, H. (1966) A maximum principle of the Pontryagin type for systems described by nonlinear difference equations. *SIAM Journal Control*, **4**, 91–111.

Hu J. and Caramanis, M. (1992) Near optimal set-up scheduling for flexible manufacturing systems, in *Proceedings of the Rensselaer Third International Conference on CIM*, Troy, NY, pp. 192–201.

Keerthy S. S. and Gilbert, E. G. (1988) Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: stability and moving-horizon approximations. *Journal of Optimization Theory and Applications*, **57**(2), 265–293.

Kimenia, J. and Gershwin, S. B. (1983) An algorithm for the control of a flexible manufacturing system. *IIE Transactions* **15**(4), 353–362.

Lasserre J. B. and Roubelat, F. (1985) Measuring decision flexibility in production planning. *IEEE Transactions on Automation and Control*. **30**(5), 447–452.

Leach, N., Makatsoris, H. and Richards, H. (1996) Supply chain control: trade-offs and systems requirements, in *Proceedings of the ESPRIT-COPERNICUS Symposium*, Budapest, Hungary, December pp. 141–154.

Liptser, R., Runggaldier, W. and Taksar, M. (1996) Deterministic approximation for stochastic control problems. *SIAM Journal of Control and Optimization*, **34**(1) 161–178.

Luh, P. B. and Hoitomt, D. J. (1993) Scheduling of manufacturing systems using the Lagrangian relaxation technique. *IEEE Transactions on Automation and Control*, **38**(7) 1066–1079.

Mayne, D. Q. and Michalska, H. (1990) Robust receding horizon control, in *Proceedings of the 29th IEEE Conference on Decision and Control*, pp. 575–581.

Pereira, F. L, Ferreira, M. M. and Sousa, J. B. (1991) An algorithm for optimal control, presented at the Second International Conference on Industrial and Applied Mathematics, Washington D. C., July.

Pereira, F. L. and Sousa, J. B. (1993) A differential inclusion approach for solving optimal control problems, in *Proceedings of the IFAC 93 World Congress*, Sydney.

Perkins, J. R. and Kumar, P. K. (1989) Stable, distributed, real time scheduling of flexible manufacturing/assembly/disassembly systems. *IEEE Transactions on Automation and Control*, **34**(2), 139–148.

Pfohl, H. (1996) Logistics. state of the art, in *Proceedings of the ESPRIT-COPERNICUS Symposium*, Budapest, Hungary, December, pp. 33–40.

Sharifnia, A. (1992) Performance of production control methods based on flow control approach, in *Proceedings of the Rensselaer Third International Conference on CIM*, Troy, NY, pp. 184–191.

Sousa, J. B. (1991) Controlo optimizado de um modelo hierárquico de processos de manufactura, M.Sc. Thesis, Porto University, Portugal.

Sousa, J. B. and Pereira, F. L. (1992) A hierarchical framework for the optimal flow control of manufacturing systems, in *Proceedings of the Renesselaer Third International Conference on CIM*, Troy, NY, pp. 278–286.

Sousa, J. B. and Pereira, F. L. (1994) A receding horizon strategy for the hierarchical control of manufacturing systems, in *Proceedings of the 4th International Conference on CIMAT, RPI*, Troy, NY, USA, Oct, pp. 443–450.

Srivatsan, N. and Gershwin, S. B. (1991) Selection of setup times in a hierarchically controlled manufacturing system, in *Proceedings of the 30th IEEE CDC*, pp. 64–49.

Vinter, R. B. and Michalska, H. (1991) Receding horizon control for nonlinear time-varying systems, in *Proceedings of the 30th IEEE CDC*, pp. 75–76.

Vinter, R. B. and Wolenski, P. R. (1994) Adjoint variables and the value function in optimal control theory: the measurable case. *Journal of Mathematical Analysis and Applications*, **53** pp. 37–51.