RAIRO. RECHERCHE OPÉRATIONNELLE

XIAOLAN XIE

Real time scheduling and routing for flexible manufacturing systems with unreliable machines

RAIRO. Recherche opérationnelle, tome 23, nº 4 (1989), p. 355-374

<http://www.numdam.org/item?id=RO_1989__23_4_355_0>

Recherche opérationnelle/Operations Research
(vol. 23, n° 4, 1989, p. 355 à 374)

REAL TIME SCHEDULING AND ROUTING FOR FLEXIBLE MANUFACTURING SYSTEMS WITH UNRELIABLE MACHINES (*)

by Xiaolan XIE [1]

Abstract. — *This paper presents a real time scheduling and routing algorithm for a failure prone Flexible Manufacturing System (FMS). The parts should be dispatched into the system at adequate times so as to minimize the disruptive effects of machine failures and to limit the work in process.*

We extend the work of Kimemia and Gershwin in which the scheduling problem is decomposed into a flow control problem and a discret part dispatching problem. In the FMSs studied here, several machines may perform one operation in different lengths of time and machines may be multi-purpose. We propose a new technique which reduces the flow control computation burden. Simulation results are presented.

Keywords : Flexible manufacturing systems; production scheduling; unreliable machines; flow control.

Résumé. — *Dans cet article, nous présentons un algorithme d'ordonnancement et de routage dans un atelier flexible dont les machines sont sujettes à pannes.*

Nous étendons le travail de Kimemia et Gershwin dans lequel le problème d'ordonnancement est décomposé en un problème de contrôle de flux de matière et un problème de lancement. Dans les ateliers flexibles étudiés dans ce papier, une opération peut être effectuée par plusieurs machines qui demandent des temps différents. En outre, une machine peut effectuer plusieurs opérations. Nous proposons une nouvelle technique qui diminue le temps de calcul du contrôle optimal de flux. Des résultats de simulation sont présentés.

Mots clés : Ateliers flexibles; ordonnancement de production; pannes de machines; contrôle de flux.

1. INTRODUCTION

Flexible manufacturing systems (FMS) are introduced to meet the requirements of a variety of part types in small and medium size. A flexible manufacturing system consists of a set of machines and some associated storage places, connected by an automated transportation system.

(*) Received December 1988.

[1] SAGEP PROJECT, INRIA-LORRAINE, Technopôle Brabois-Nancy, Campus Scientifique, boulevard des Aiguillettes, B.P. n° 239, 54506 Vandœuvre-les-Nancy, France.

Recherche opérationnelle/Operations Research, 0399-0559/89/04 355 20/$ 4.00

A FMS system is often designed to be able to produce a mixed set of parts simultaneously. Multi-purpose machines are introduced to meet this need. On the other hand, multiple production facilities which can perform the same operations may exist. This allows the production to continue even if some machines fail to work.

However, the control of a FMS is very difficult because of its complexity. It is well known that the control of a FMS is NP-hard. In addition, the fact that the system is subject to many random disruptive events, such as machine breakdowns, material unavailability, etc., makes control even more difficult.

Scheduling problems have been addressed by many authors. A complete survey can be found in Graves [1981]. Most authors state the problem as a mixed integer programming problem. This approach leads to large scale problems, even for small production systems.

Hierarchical scheduling algorithms have been proposed by Hildebrant [1980], Kimemia and Gershwin [1983], Gershwin, Akella and Choong [1985], Gershwin [1987 *a*, *b*]. These algorithms avoid combinatorial features. It seems to be a promising way to do real time production scheduling.

In this paper, we assume that the production requirements are specified at a higher level of the hierarchy, *i. e.* the master planning level. The control problem is two-fold. First, the parts must be loaded into the system at a certain rate in order to meet the production requirements. Second, the parts must be routed correctly in the system so that no congestion should occur and the work in process is as small as possible.

The following assumptions are made in this paper.

1. The time required for a machine to switch from one operation to another one is small as compared to the part processing times.

2. The part processing times are small compared to the mean times between failures and the mean times to repair.

3. The mean times between failures and the mean times to repair are great as compared to the processing times.

4. The short term scheduling horizon is greater than the mean times between failures and the mean times to repair.

For systems of this type, Kimemia and Gershwin [1983] propose a two level controller. At the higher level, they use a continuous representation of the material flow. The failures and the repairs are modelled by Markov processes. The control problem consists in finding a production rate so as to keep the production close to the production requirements. It is stated as an optimal control problem for a system with jump Markov disturbances, which has been solved from control theory by Rishel [1975]. At the lower level, the pairs are dispatched into the system according to the production rate specified at the upper level. The main advantage of this approach is that the scheduling becomes very easy once the production rate is known.

An important improvement of this approach has been made by Gershwin, Akella and Choong [1985] for systems without parallel machines. They use a quadratic approximation of the cost functions which are obtained by solving a complex Bellman's equation in Kimemia and Gershwin's algorithm. A technique to compute parameters of these cost functions is proposed. This algorithm greatly reduces the computational burden and avoids the control singularity.

Recently, Maimon and Gershwin [1988] have also proposed a control algorithm for systems with non-identical parallel machines. Simulation results are not presented.

Akella and Kumar [1986], Bielecki and Kumar [1988], and Sharifnia [1988] obtain analytic solution to one product problem. Bielecki and Kumar point out that a zero-inventory policy can be optimal even in the presence of uncertainty.

Gershwin [1987 *a*, *b*] proposes a hierarchical control framework to deal with a much richer catalog of events, including setups, machine failures, preventative maintenance, etc. The basic idea is to treat events of different frequencies separately.

The purpose of this paper is to describe a new extension of the above mentioned work for systems with non-identical parallel machines. Quadratic approximation of the cost functions is used. A new technique is proposed to compute parameters of these quadratic functions.

Outline of paper

The paper is organized as follows. In Section 2, the FMS model is presented. Characteristics of optimum control policy and optimum control computation techniques are reviewed and a control framework is proposed. In Section 3, we propose an approximation of the cost functions. In Section 4, the discrete scheduling problems, including part dispatching and the choice of parts for idle machines, are addressed. In Section 5, the long term production capacity is examined. Simulation results are given in Section 6, and conclusions are given in Section 7.

2. FLOW CONTROL MODEL

In this section, the FMS model and the flow control problem are described. Then, the characteristics of optimum control policies of the problem and the computation techniques of the optimum control are reviewed. We finally propose a production control framework.

2.1. Problem formulation

The system under study consists of M machines on which N part types are produced. The manufacturing process of a part type can be described by a sequence of operations. Each operation may be done on several machines. Let $u(t)$ be the production rate vector of part types at time t. Let d_n be the specified demand rate for type n parts.

Let $x_n(t)$ be the difference between the total production and the cumulative demand of type n parts up to instant t. It is given by

$$dx_n(t)/dt = u_n(t) - d_n. \tag{1}$$

A machine is either up or down. A boolean variable, denoted by $\alpha_m(t)$, is used to indicate the state of machine m,

$$\alpha_m = \begin{Bmatrix} 1 & \text{machine } m \text{ is up;} \\ 0 & \text{otherwise.} \end{Bmatrix} \tag{2}$$

The time between failures and the time to repair are assumed to be exponentially distributed random variables.

The machine state of the system, denoted by $\alpha(t)$, is defined by

$$\alpha(t) = (\alpha_1(t), \alpha_2(t), \ldots, \alpha_M(t)).$$

It can be modelled by a Markov chain. Let S be the set of possible machine states of the system, then for any α, $\beta \in S$ and $\alpha \neq \beta$

$$P[\alpha(t+\delta t) = \beta / \alpha(t) = \alpha] = \lambda_{\alpha\beta}\, \delta t \tag{3}$$

where $\lambda_{\alpha\beta}$ are functions of the failure rates and the repair rates of machines.

The production rate is limited by the rate at which machines can do operations. No machines can work more than 100% of the time and machines under repair cannot work at all. Let y_{nm}^k be the rate at which type n parts pass through machine m for operation k, and let τ_{nm}^k be the time required to complete the operation. Then, the rate y_{nm}^k must satisfy the following constraints

$$\sum_n \sum_k \tau_{nm}^k y_{nm}^k \leqq \alpha_m, \qquad \forall m. \tag{4}$$

It is assumed that no material is accumulated within the system. Consequently, the rate at which type n parts are sent for operation k is equal to the production rate $u_n(t)$. This is expressed as

$$\sum_m y_{nm}^k = u_n(t), \qquad \forall n, k. \tag{5}$$

Let $\Omega(t)$ denote the set of $u(t) = (u_1(t), \ldots, u_N(t))$ such that there exist flow rates y_{nm}^k satisfying equations (4) and (5). It is function of the machine state of the system and it is a convex polyhedron. Then, it can be written as $\Omega[\alpha(t)]$.

A production rate $u(t)$ is feasible if and only if

$$u(t) \in \Omega[\alpha(t)].$$

The flow control problem can be stated as follows. Given a initial buffer state $x(t_0)$ and machine state $\alpha(t_0)$, find a control policy $u(t) \in \Omega[\alpha(t)]$ for $t_0 \leqq t \leqq T$ that minimizes the following performance index

$$J_u(x_0, \alpha_0, t_0) = E\left\{ \int_0^T [g(x(t))\, dt / x(t_0) = x_0, \alpha(t_0) = \alpha_0] \right\}. \tag{6}$$

subject to (1), (3). The function $g[x(t)]$ is a positive convex function which penalizes the controller for failing to meet the demand and for keeping inventory.

2.2. Characteristics of optimum control

The control problem has been largely studied by Rishel [1975]. Kimemia and Gershwin [1983]. It is shown that for any given initial buffer level x and machine state α, the optimum policy $u^*(t)$ can be determined by

$$\min_{u \in \Omega(\alpha)} \frac{\partial}{\partial x} J_{u^*}(x, \alpha, t)\, u \tag{7}$$

where $J_{u^*}(x, \alpha, t)$ is the expected çost, when the control policy u^* is applied in the interval $[t, \mathrm{T}]$.

The expression (7) is linear in u and $\Omega(\alpha)$ is a convex polyhedral set. An optimal policy $u^*(x, \alpha, t)$ takes values at extreme points of $\Omega(\alpha)$ whenever the gradient $\partial/\partial x\, J_{u^*}(x, \alpha, t)$ exists. For each machine state α, an optimal policy divides the buffer state space into a set of regions in which the production rate is constant. But, the gradient $\partial/\partial x\, J_{u^*}(x, \alpha, t)$ does not always exist at the boundaries of these regions.

The cost function has the following properties:

1. $J_{u^*}(x, \alpha, t)$ is a continuous function of x and t for all $\alpha \in S$;
2. $J_{u^*}(x, \alpha, t)$ is a convex function of x for $t \in (0, T)$ and $\alpha \in S$.
3. $\lim\limits_{|x| \to \infty} J_{u^*}(x, \alpha, t) = \infty$.
4. There exists a finite x that minimizes $J_{u^*}(x, \alpha, t)$.

In the time invariant case for which $d(t) = d$ and $T = \infty$, J_u is the average cost rather than a total cost over the interval $[t, T]$. As a result, the cost function $J_u(x(t), \alpha(t), t)$ does not depend explicitly on t and can be written as $J_u(x(t), \alpha(t))$.

There are two kinds of machine states: those for which the demand rate d is feasible, *i. e.* those for which $d \in \Omega(\alpha)$, and those for which the demand rate d is infeasible.

For any feasible state α, if the system remains in state α for a sufficient period of time, the production rate will become equal to the demand rate and the buffer level will remain constant. This fixed buffer level, noted by x_α^H, will be called the hedging point. It is a buffer level hedged against the future breakdowns. This can be explained by the following equation

$$\frac{dJ_{u^*}(x(t), \alpha(t))}{dt} = \frac{\partial J_{u^*}(x(t), \alpha(t))}{\partial x}(u^* - d).$$

Since u^* minimizes $\partial/\partial x, J_{u^*}(x, \alpha)\, u$ for all $u \in \Omega(\alpha)$, and since $d \in \Omega(\alpha)$, then $d/dt\, J_{u^*}(x(t), \alpha(t))$ is negative and $J_{u^*}(x(t), \alpha(t))$ is a decreasing function of time t. On the other hand, J_{u^*} is a positive quantity. $J_{u^*}(x(t), \alpha(t))$ decreases to a limit. The limit is reached when $x(t)$ minimizes $J_{u^*}(x(t), \alpha(t))$. After that time, u^* is equal to the demand rate and the buffer level remains constant. The x that minimizes $J_{u^*}(x(t), \alpha(t))$ is the hedging point.

2.3. Optimal control computation techniques

As a result of the above discussion, the optimum control can be computed by a two step procedure. At the first step, the cost function is evaluated. This consists in finding out the optimum law of feedback. At the second step, the production rate is computed by solving the linear problem (7).

The main difficulty of the optimum control computation lies in cost function evaluation. The exact solution consists in solving a partial differential equation.

Kimemia [1983] observed an undesirable singular behavior when applying this approach. This occurs when the buffer level reaches a boundary of a region in which the control is constant and it should remain at this boundary.

To overcome this difficulty, Gershwin, Akella and Choong [1985] propose a quadratic approximation of cost functions. With this approximation, the optimum control policy becomes piece-wise constant. For any initial machine state and buffer level, the optimum control policy can be described by a sequence of controls and their application times $\{(u^0, t^0), \ldots, (u^{\mathcal{N}-1}, t^{\mathcal{N}-1}), (u^{\mathcal{N}}, t^{\mathcal{N}} = \infty)\}$ if the machine state lasts for a sufficient period of time. Another sequence of controls will not be applied until the machine state has changed. Notice that these sequences can be computed easily. Control singularity is avoided when applying this approach.

This approach works very well for all our simulation experiences. This is because the exact definition of function $g(.)$ is not important, what matters is to have the system behave property. A quadratic approximation of the cost function is capable of giving an adequate feedback of the buffer state. For systems without parallel machines, Gershwin, Akella and Choong [1985] propose an adequate technique to compute the parameters of the cost functions.

2.4. Proposed control framework

Figure 2.1 summarizes our control framework. First, the cost functions are estimated. As we use a quadratic approximation, this step consists in computing parameters of these quadratic functions. Second, the production rate $u(t)$ is computed by equation (7). Then, routing rates $y_{nm}^k(t)$ are determined from $u(t)$ by balancing the machine workloads. At the discrete scheduling level, parts are loaded into the system according to production rate $u(t)$ computed above and are sent to idle machines according to routing rates $y_{nm}^k(t)$.

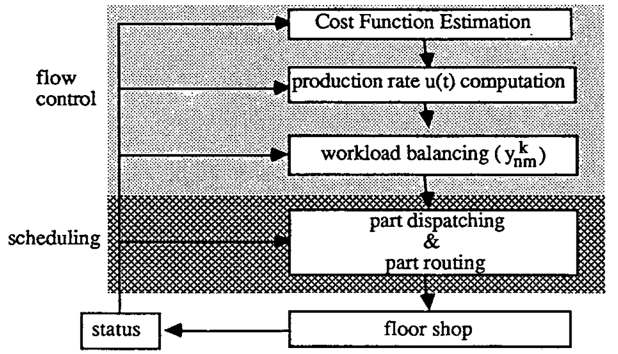

Figure 2.1. — Control computation framework.

3. COST FUNCTION ESTIMATION

As pointed out above, what matters is not the details of cost function, but a proper system behaviour. The expected system behaviour is analyzed in this section. Then, we propose a quadratic cost function approximation of cost functions and a technique to compute parameters of these quadratic functions.

3.1. Expected system behaviour

For a feasible system state α, *i.e.* $d \in \Omega(\alpha)$, Figure 3.1 (*a*) demonstrates a typical trajectory of the buffer level $x_n(t)$. The system state α becomes feasible at time t_0. The production rate is chosen to be greater than the demand rate. Then, the buffer level moves toward the hedging poont $H_n(\alpha)$. When the

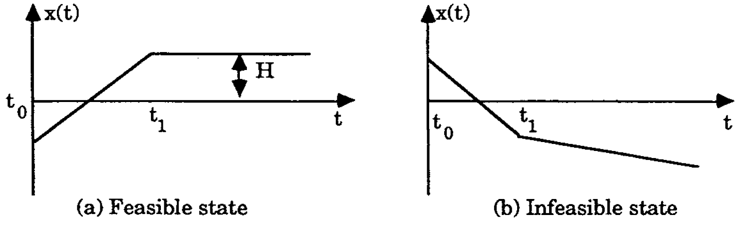

Figure 3.1. — Typical buffer trajectory.

hedging point has been reached, the production rate is equal to the demand rate and the buffer level remains at hedging point.

For any infeasible state α, *i. e.* $d \notin \Omega(\alpha)$, Figure 3.1 (*b*) shows a typical trajectory of $x_n(t)$. The system state α becomes infeasible at time t_0. The production cannot keep up with the demand. Then, the buffer level decreases. If the system state lasts, the production rate tends to become constant and the buffer level tends to decrease at a fixed rate.

For any state, the production rate becomes constant in finite time if the system state endures. Let $u^{\mathcal{N}}(\alpha, x_0)$ denote this constant if the system state is α and the buffer level is x_0 when the system state becomes α. It has the following properties:

(1) $u^{\mathcal{N}}(\alpha, x_0) \leqq d$ and $u^{\mathcal{N}}(\alpha, x_0) \in \Omega(\alpha)$;

(2) if the system state is feasible, *i. e.* $d \in \Omega(\alpha)$, then $u^{\mathcal{N}}(\alpha, x_0) = d$.

We introduce $d^c(\alpha)$, the *controllable demand rate.* It is determined by

$$d^c(\alpha) = \underset{u \in \Omega(\alpha) \text{ and } u \leqq d}{\text{Argmin}} \sum_n b_n^* (d_n - u_n)$$

where b_n is the cost incurred for backlogging one type n part.

For any feasible state, *i. e.* $d \in \Omega(\alpha)$, the controllable demand is equal to the demand, *i. e.*

$$d^c(\alpha) = d.$$

For any infeasible state α, the mean incurred cost is minimized if the state lasts and $u^{\mathcal{N}}(\alpha, x_0)$ is equal to $d^c(\alpha)$. So, the controllable demand rate may be seen as a desirable production rate if the system state lasts.

3.2. Cost function estimation

When the system state becomes infeasible, the production cannot keep up with the demand. We propose a quadratic approximation of the cost function, which leads the production rate toward the controllable demand rate if the state lasts, *i. e.* $u^{\mathcal{N}}(\alpha, x_0) = d^c(\alpha)$.

We introduce $x_n^c(t)$, a fictitious buffer level. It is determined by

$$\left.\begin{aligned} \frac{dx_n^c(t)}{dt} &= u_n(t) - d_n^c(\alpha(t)) \quad \text{if} \quad \alpha(t) = \alpha(t-0); \\ x_n^c(t) &= x_n(t) - \mathrm{H}_n^u(\alpha(t)) \quad \text{if} \quad \alpha(t) \neq \alpha(t-0). \end{aligned}\right\} \tag{9}$$

To understand this quantity, imagine that the demand rate d_n is split into two part $d_n^c(\alpha(t))$ and $d_n - d_n^c(\alpha(t))$. The rate $d_n - d_n^c(\alpha(t))$ represents the lost demand rate, but we assume that a buffer level $H_n^u(\alpha(t))$ has been hedged against this lost demand. Thus, $x_n^c(t)$ can be seen as the buffer level for the controllable demand.

We use the following quadratic approximations of the cost functions:

$$J(x, \alpha) = \frac{1}{2}(x^c(t) - H^\alpha)^T A(\alpha)(x^c(t) - H^\alpha) + c(\alpha) \tag{10}$$

where $A(\alpha)$ is a positive definite diagonal matrix, $c(\alpha)$ is a vector of positive components and H^α is the hedging point for the controllable demand.

This is an extension of the quadratic approximation proposed by Gershwin, Akella and Choong [1985]. It is easy to show that if a system state lasts, the final production rate becomes equal to the controllable demand rate.

3.3. Hedging point computation

In this section, we propose a myopic technique to compute hedging points. For each state, only the failures of machines up in the state are taken into account.

Let us consider a system state α and a machine m up in state α, *i. e.* $\alpha_m = 1$. Figure 3.2 shows a typical trajectory of the buffer level $x_n^c(t)$ when machine

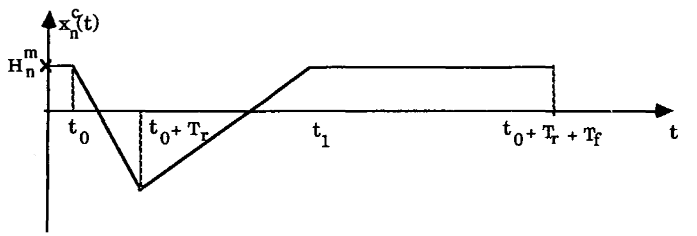

Figure 3.2. — Typical buffer level trajectory.

m breaks down. Machine m breaks down at time t_0; the buffer level decreases at rate $-\Delta d_n$; the failure lasts for a length of time T_r. After the repair, the production rate is assigned a value U_n greater than the controllable demand rate $d_n^c(\alpha)$. The buffer level moves toward the hedging point. When the hedging point is reached, the buffer level stays at the hedging point and the production rate becomes equal to the controllable demand rate.

For this trajectory, we can evaluate the cost by penalyzing the positive area with weight a_n and the negative area with weight b_n. By minimizing the total weighted area, we have

$$H_n^m = \frac{b_n}{a_n + b_n} T_r \Delta d_n - \frac{a_n d_n^c(\alpha)}{(a_n + b_n) U_n} (T_f (U_n - \Delta d_n) - T_r \Delta d_n) \tag{11}$$

If the following equation is true

$$T_f (U_n - \Delta d_n) = T_r \Delta d_n, \tag{12}$$

the computation can be further simplified as follows

$$H_n^m = b_n / (a_n + b_n) T_r \Delta d_n. \tag{13}$$

Notice that equation (12) means that the system has just enough time to recover the demand lost during the failure.

Assuming the controllable demand to be a new demand, the new controllable demand d^{c^*} is computed when machine m fails. The rate Δd is determined by

$$\Delta d = d^c(\alpha) - d^{c^*}. \tag{14}$$

The above computed H_n^m is the buffer hedged against the failure of machine m. We propose the following value for the hedging point H_n^α

$$H_n^\alpha = \max_{\alpha_m = 1} H_n^m(\alpha). \tag{15}$$

3.4. Matrix A (α) computation

In the previous work, little attention has been paid to the computation of the weighting matrix $A(\alpha)$. Nevertheless, simulation results reveal that it is an important factor. The computation of the matrix must include the vulnerability of a part to failures, the capacity of the system to recover the lost demand and the relative importance of the part.

Matrix A is a positive definite diagonal matrix. The value of A_{jj} reflects the relative priority of part type j. We propose the following value for A_{jj}

$$A_{jj} = \beta_j * \gamma_j * \xi_j. \tag{16}$$

where β_j, γ_j, ξ_j measure respectively the relative importance of type j parts, the vulnerability of type j parts to failures, and the difficulty of recovering the lost demand for type j parts.

As for the vulnerability of parts to failure, we propose the following value

$$\gamma_j = \sum_{k \in K_j} \sum_m \gamma_{nm}^k (\text{MTTR}_m / \text{MTBF}_m) \tag{17}$$

where MTTR_m and MTBF_m denote the mean time to repair and mean time between failures of machine m, and K_j is the set of operations for type j parts.

γ_{nm}^k is the proportion of type n parts to be sent to machine m for operation k if the system is in steady state. It is computed by balancing the machine workloads ($\text{MTTR}_m/\text{MTBF}_m$) reflects the reliability of the mahcine. The summation in equation (17) reflects the vulnerability of the part to failures. The more operations a part requires, the more vulnerable the part is to failures. The smaller the mean time to repair and the greater the mean time between failures, the more vulnerable the part is to failures.

For the difficulty of recovering lost demand, the following value is used

$$\xi_j = 1/U_j^* \tag{18}$$

where U_j^* is the maximal rate at which the system can produce type j parts if the system is in steady state. The greater U_j^* is, the more easily the system can recover the lost demand for type j parts during a failure.

4. LOWER LEVEL SCHEDULING

There are two problems at the discrete scheduling level. One problem is to determine the instants at which parts are dispatched into the system. The other problem is to load parts on idle machines, *i. e.* the routing problem.

As for the part dispatching, the rule proposed by Gershwin, Akella and Choong [1985] is used. Let $x_n^A(t)$ be the actual production surplus of the part type n defined as follows

$$x_n^A(t) = [\text{number of parts of type } n \text{ loaded during } (0, t)] - d_n t. \tag{19}$$

The strategy consists in trying to keep $x_n^A(t)$ always greater than $x_n(t)$ and of loading a part of type n every time $x_n^A(t)$ is smaller than $x_n(t)$, *i. e.*

$$x_n^A(t) < x_n(t),$$

and not loading otherwise.

For the routing problem, define $x_{nm}^{A}(t)$ and $x_{nm}^{k}(t)$ as follows

$$x_{nm}^{A\,k}(t) = [\text{number of type } n \text{ parts loaded on machine } m \text{ for operation } k \text{ during } (0, t)]; \qquad (21)$$

$$x_{nm}^{k}(t) = \int_0^t y_{nm}^{k}(s)\,ds - x_{nm}^{A\,k}(t).$$

x_{nm}^{k} measures the cumulative difference between the number of parts assigned to machine m for operation k at the flow control level and the real number of parts assigned to machine m for operation k. Let $\mathbf{q}_{\mathbf{n}}^{\mathbf{k}}$ be the number of type n parts waiting for operation k. The following rule is used to choose a part to be loaded on an idle machine from parts waiting in the corresponding queue.

$$\min_{n,\,k,\,m} \{ x_{nm}^{k}(t)/q_n^k(t) > 0 \text{ and machine } m \text{ is idle} \}. \qquad (21)$$

Let (n^*, k^*, m^*) be an optimal solution. Then, a type n^* part waiting for operation k^* is loaded on machine m^*.

5. LONG TERM CAPACITY DISCUSSION

In this section, we are interested in the total production capacity of the system over the short term horizon as opposed to the instantaneous capacity discussed in Section 2.

Let $p(\alpha)$ denote the probability that the system state is α, and let $u(\alpha)$ be the production rate vector when the system state is α. Assume that a production ratio γ_n is assigned to each part type. Then, the maximal throughput of the system with fixed ratios, denoted by $X^*(\gamma_1, \ldots, \gamma_n, \ldots, \gamma_N)$, can be computed as follows

$$X^*(\gamma_1, \ldots, \gamma_n, \ldots, \gamma_N) = \max X \qquad (22)$$

subject to

$$\gamma_n X = \sum_{\alpha} p(\alpha)\, u_n(\alpha), \qquad \forall n;$$

$$u(\alpha) \in \Omega(\alpha), \qquad \forall \alpha \in S.$$

Let Ω^L denote the set of feasible controls in long run. A long term production rate vector is said to be feasible, *i. e.* $u \in \Omega^L$, if and only if

$$\sum_j u_j \leqq X^*(\gamma_1, \ldots, \gamma_n, \ldots, \gamma_N).$$

Since all the sets $\Omega(\alpha)$ are convex polyhedral and the long term capacity set Ω^L is a linear projection of these sets, Ω^L is also a convex polyhedral set.

6. SIMULATION RESULTS

An example and simulation results are presented in this section. The performance of the algorithm is discussed.

A flexible transfert line is shown in Figure 6.1. It consists of three workstations. Each workstation consists of two machines of different performance. Two types of parts are produced. The first part type requires three

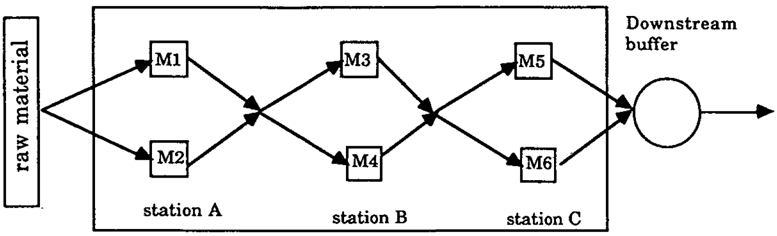

Figure 6.1. — A flexible transfer line.

operations, one at each workstation; the second part type requires two operations, one at the first workstation and another one at the third workstation. The demand is 1 part per minute for type 1 parts, and 2 parts per minute for type 2 parts.

The processing times were given in Table I. The ability data including the mean time between failures (MTBF) and the mean time to repair (MTTR) are given in Table II. The expected utilization and availability of the machines are given in Table III. By utilization ratio of a machine we mean the ratio between the machine busy time and the machine up time. Station *A* and station *C* are heavily loaded.

The simulation model was run for 7,200 minutes (120 hours). Table IV shows the real machine availability and real machine utilization. As expected,

TABLE I

Processing times in minutes.

part	M1	M2	M3	M4	M5	M6
1	1	1	1	3	1	1
2	1/2	1/3	Not required		1/3	1/3

TABLE II

Reliability data in minutes.

	M1	M2	M3	M4	M5	M6
MTBF	100	200	200	100	300	100
MTTR	10	10	20	10	30	10

TABLE III

Expected utilization and availability.

	M1	M2	M3	M4	M5	M6
Availability	0.909	0.952	0.909	0.909	0.909	0.909
Utilization	0.895	0.895	0.825	0.825	0.917	0.917

stations *A* and *C* are heavily loaded. The machine workloads are well balanced, 90.69 and 86.91% for the two machines of the station *A*, 81.37 and 83.02% for those of station *B*, and 90.46 and 90.76% for those of station *C*.

Table V gives the production statistics. On the average, the production was 54.1 pieces behind demand for type 1 parts and 119.5 pieces behind demand for type 2 parts. The average in-process inventory in the system was small, 5.199 type 1 pieces and 3.274 type 2 pieces. At the end of the simulation, the number of parts actually produced was 7,053 (or 98.0%) and 13,984 (or 97.1%). The average over-production was 0.606 type 1 pieces and 0.428 type 2 pieces. The average backlog was 54.7 type 1 pieces and 119.9 type 2 pieces.

TABLE IV
Real utilization and availability.

	M1	M2	M3	M4	M5	M6
Availability	0.930	0.940	0.889	0.926	0.891	0.905
Utilization	0.907	0.869	0.814	0.830	0.905	0.908

TABLE V
Production statistics.

Part	Demand	Production	Mean WIP	Mean buffer	Mean Surplus	Mean Backlog
1	7200	7053	5.199	-54.12	0.606	54.73
2	14400	13984	3.274	-119.5	0.428	119.9

TABLE VI
A sample production trajectory.

part	type1			type2		
instant (minutes)	demand	production	%_produced	demand	production	%_produced
0	0	0	-	0	0	-
600	600	582	97.0	1200	1146	95.5
1200	1200	1176	98.0	2400	2385	99.4
1800	1800	1794	99.7	3600	3598	99.9
2400	2400	2371	98.8	4800	4723	98.4
3000	3000	2951	98.4	6000	5964	99.4
3600	3600	3588	99.7	7200	7135	99.1
4200	4200	4160	99.0	8400	8318	99.0
4800	4800	4749	98.9	9600	9453	98.5
5400	5400	5313	98.5	10800	10798	100.0
6000	6000	5829	97.1	12000	11597	96.6
6600	6600	6421	97.3	13200	12753	96.6
7200	7200	7053	98.0	14400	13984	97.1

Table VI shows a sample trajectory of the simulation. The production may leave from the demand in case of long machine failures. In any case, the controller is able to keep the production ratio close to the required ratio. That is an important feature of this controller.

TABLE VII

Sojourn times of system states over 7,200 minute simulation.

$\alpha_6\alpha_5\alpha_4\alpha_3\alpha_2\alpha_1$	duration (minutes)	$\alpha_6\alpha_5\alpha_4\alpha_3\alpha_2\alpha_1$	duration (minutes)
000000	0	100000	0
000001	0	100001	0
000010	0	100010	0
000011	0	100011	0
000100	0	100100	0
000101	0	100101	0
000110	0	100110	0
000111	3	100111	68
001000	0	101000	0
001001	0	101001	0
001010	0	101010	6
001011	0	101011	63
001100	0	101100	0
001101	7	101101	33
001110	29	101110	8
001111	87	101111	481
010000	0	110000	0
010001	0	110001	8
010010	0	110010	0
010011	2	110011	49
010100	0	110100	0
010101	5	110101	21
010110	0	110110	58
010111	44	110111	274
011000	0	111000	0
011001	0	111001	51
011010	0	111010	10
011011	26	111011	583
011100	1	111100	3
011101	36	111101	266
011110	26	111110	355
011111	414	111111	4178

System capacity and performance of the controller

Given a simulation experience ω run over time length $T(\omega)$, the time lengths for which the system stays in each particular state, denoted by $T(\alpha, \omega)$, are known. We are interested here in the system capacity for such a particular simulation experience.

Let $X^*(\gamma_1, \ldots, \gamma_N; \omega)$ denote the maximal throughput with fixed production ratio of the simulation experience ω. Its computation is the same as that of X^* in Section 6 by replacing $p(\alpha)$ by $T(\alpha, \omega)/T(\omega)$.

Table VII shows the system's sejourn times in each state of a simulation experience run over 7,200 minutes. The maximal throughput of the simulation experience is

$$X^*(\gamma_1, \ldots, \gamma_N; \omega) = 2.975 \text{ parts/minute.}$$

More precisely, the production capacity is 7,140 type 1 parts (99.2%) and 14,280 type 2 parts (99.2%). The demand cannot be met.

We remind that the real production is 7,053 type 1 parts (98.0%) and 13,984 type 2 parts (97.1%). This shows how performant the controller is.

Justification of Parameters of cost functions

In Section 3.4, we have proposed the following value for the weighting factor of cost functions A_{jj}

$$A_{jj} = \beta_j * \gamma_j * \xi_j.$$

Table VIII shows the simulation result with $\xi_j = 1$. The percentage of type 2 parts produced is always greater than that of type 1 parts. Compared with the result in Table VI, the production ratio is far from the demand ratio. But, the total productions in the two simulations are very close. A FMS may be one part of whole manufacturing system, parts produced by the FMS may be assembled later. In this respect, the result in Table VI is better than that in Table VIII.

7. CONCLUSION

This paper extends the work of Gershwin and his co-workers on FMS's hierarchical production scheduling. Multi-purpose machines and parallel machines of different performance are taken into account.

TABLE VIII
Simulation result with $\xi_j = 1$.

part	type1			type2		
instant (minutes)	demand	production	%_produced	demand	production	%_produced
0	0	0	-	0	0	-
600	600	566	94.3	1200	1181	98.4
1200	1200	1163	96.9	2400	2398	99.9
1800	1800	1793	99.6	3600	3598	99.9
2400	2400	2354	98.1	4800	4772	99.4
3000	3000	2923	97.4	6000	5979	99.6
3600	3600	3558	98.8	7200	7153	99.3
4200	4200	4103	97.7	8400	8339	99.3
4800	4800	4673	97.4	9600	9501	99.0
5400	5400	5241	97.1	10800	10800	100.0
6000	6000	5686	94.8	12000	11807	98.4
6600	6600	6270	95.0	13200	12987	98.4
7200	7200	6918	96.1	14400	14172	98.4

The basic idea of the hierarchical production scheduling approach is to use a continuous representation of the material flow and to guide the discrete scheduling by the results of optimal flow control.

We introduce the controllable demand rate at the flow control level. This allows as to propose a new hedging point computation technique. More attention is paid to the computation of quadratic function parameters.

As for the discrete scheduling level, some very simple rules are used to dispatch parts into the system and to load parts on idle machines by using the decisions made at the flow level. Simulation results shows that the algorithm works very satisfactorily.

Further research work consist in including disruptive events other than machine failures for instance machine set-ups. There are two type of events, controllable or uncontrollable. The main difficulty lies in the wide variety of event occurrence frequencies and the scheduling of controllable events.

REFERENCES

1. R. AKELLA, Y. F. CHOONG and S. B. GERSHWIN, *Performance of Hierarchical Production Scheduling Policy*, IEEE Trans. On Components, Hybrids, and Manufacturing Technology, Vol. CHMT-7, No. 3, September 1984.

2. R. Akella and P. R. Kumar, *Optimal Control of Production Rate in a Failure Prone Manufacturing System*, IEEE Trans. on Automatic Control, Vol. AC-31, No. 2, February 1988.
3. T. Bielecki and P. R. Kumar, *Optimality of Zero-Inventory Policies for Unreliable Manufacturing Systems*, Operations Research, Vol. 36, No. 4, July-August 1988.
4. S. B. Gershwin, *A Hierarchical Framework for Discret Event Scheduling in Manufacturing Systems*, Presented at IIASA Workshop on Discret Event Systems: Models and Applications, Sopron, Hungary, August 3-7, 1987. Published in Vol. 103, Lecture Notes in Control and Information Sciences, Discret Event Systems: Models and Applications, P. Varaiya and A. B. Kurzhanski Eds., Springer-Verlag, 1987 *a*.
5. S. B. Gershwin, *A Hierarchical Framework for Manufacturing Systems Scheduling: A Two-Machine Example*, Proceedings of the 26th IEEE Conference on Decision and Control. Los Angeles, California, December 1987 *b*.
6. S. B. Gershwin, R. Akella, and Y. F. Choong, *Short-term Production Scheduling of an Automated Manufacturing Facility*, IBM Journal of Research and Development, Vol. 29, No. 4, July 1985.
7. S. B. Gershwin, R. R. Hildebrandt, R. Suri and S. K. Mitter, A Control Theorist's Perspective on Recent Trends in Manufacturing System, IEEE Control Systems Magazine, Vol. 6, No. 2, April 1986.
8. S. C. Graves, *A Review of Production Scheduling*, Operations Research, Vol. 29, No. 4, July-August 1981.
9. R. R. Hildebrandt, *Scheduling and Control of Flexible Machining Systems when Machines are Prone to Failures*, Ph. D. Thesis, M.I.T. Dept. of Astronautics and Aeronautics, August 1980.
10. J. Kimemia, *Hierarchical Control of Production in Flexible Manufacturing Systems*, Ph. D. Thesis, M.I.T. Dept of Electrical Engineering and Computer Science, April 1982.
11. J. Kimemia and S. B. Gershwin, *An Algorithm for the Computer Control of a Flexible Manufacturing System*, IIE Trans., Vol. 15, No. 4, December 1983.
12. O. Z. Maimon and S. B. Gershwin, *Dynamic Scheduling and Routing For Flexible Manufacturing Systems that Have Unreliable Machines*, Operation Research, Vol. 36, No. 2, March-April 1988.
13. R. Rishel, *Dynamic Programming and Minimum Principes for Systems with Jump Markov Disturbances*, SIAM Journal on Control, Vol. 13, No. 2, February 1975.
14. A. Sharifnia, *Production Control of a Manufacturing System with Multiple Machine States*, IEEE Trans. on Automatic Control, Vol. AC-33, No. 7, July 1988.