A Control Perspective on Recent Trends in Manufacturing Systems

Stanley B. Gershwin, Richard R. Hildebrant, Rajan Suri,
and Sanjoy K. Mitter

INTRODUCTION

While the technology of manufacturing (including processes and computer hardware and software) is improving rapidly, a basic understanding of the *systems issues* remains incomplete. These issues include production planning, scheduling, and control of work in process. They are complicated by randomness in the manufacturing environment (particularly due to machine failures and uncertainty and variability in production requirements), large data requirements, multiple-level hierarchies, and other issues that control engineers and systems engineers have studied in other contexts.

The purpose of this paper is to present an interpretation of recent progress in manufacturing systems from the perspective of control. We believe that this community has a vocabulary and a view of systems that can be helpful in this area. However, in order for this group to make that impact, it is essential that they learn the problems and terminology and become familiar with recent research directions. This paper is intended to present certain issues in manufacturing management in a way that will facilitate this.

We will establish a framework for manufacturing systems issues that is heavily influenced by control and systems thinking. We will then summarize current practice and current research, and critique them from the point of view of that framework.

GENERAL PERSPECTIVE

The purpose of manufacturing system control is not different in essence from many other control problems: it is to ensure that a complex system behave in a desirable way. Many notions from control theory are relevant here, although their specific realization is quite different from more traditional application areas. The standard control theory techniques do not apply: we have not yet seen a manufacturing system that can be usefully represented by a linear system with quadratic objectives. This is not surprising; standard techniques have been developed for what have been standard problems. Manufacturing systems can be an important area for the future of control; new standard techniques will be developed.

Some central issues in manufacturing systems include *complexity, hierarchy, discipline, capacity, uncertainty,* and *feedback.* Important notions of control theory include state and control variables, the objective function, the dynamics or plant model, and constraints. It would be premature to try to identify these with all the issues outlined in this paper; it would even go against the purpose of the paper, which is to stimulate such modeling activity. In this section, we describe the relevance of these notions to the manufacturing context for readers whose primary background is in control and systems.

Complexity

Manufacturing systems are large-scale systems. Enormous volumes of data are required to describe them. Optimization is impossible; suboptimal strategies for planning based on hierarchical decomposition are the only ones that have any hope of being practical.

Hierarchy

There are many time scales over which planning and scheduling decisions must be made. The longest term decisions involve capital expenditure or redeployment. The shortest involve the times to load individual parts, or even robot arm trajectories. While these decisions are made separately, they are related. In particular, each long-term decision presents an assignment to the next shorter term decision maker. The decision must be made in a way that takes the resources — i.e., the capacity — explicitly into account. The definition of the capacity depends on the time scale. For example, short time-scale capacity is a function of the set of machines operational at any instant. Long time-scale capacity is an average of short time-scale capacity.

Machine-Level Control

At the very shortest time scale is the machine-level control. This includes the calculation and implementation of optimal robot arm trajectories; the design of "ladder diagrams" for relays, microswitches, motors, and hydraulics in machine tools; and the control of furnaces and other steps in the fabrication of semiconductors. Other short scale issues include the detailed control of a cutting tool: in particular, adaptive machining.

There is no rule that determines exactly what this shortest time scale is. A robot arm movement can take seconds while a semiconductor oxidation step can take hours.

The issue at this time scale is the optimization of each individual operation. Here, one can focus on minimizing the time or other cost of each separate movement or transformation of material. One can also treat the detailed relationships among operations. An example of this is the *line balancing* problem. Here, a large set of operations is grouped into tasks to be performed at stations along a production line. The objective is to minimize the maximum time at a station, which results in maximum production rate.

Other control problems at this time scale include the detection of wear and breakage of machine tools, the control of temperatures and partial pressures in furnaces, the automatic control of the insertion of electronic components into printed circuit boards, and a vast variety of others.

Cell Level

At the next time scale, one must consider the interactions of a small number of machines. This is cell-level control and includes the operation of small, flexible manufacturing systems. The important issues include routing and scheduling. The control problem is ensuring that the specified volumes are actually produced. At this level, the detailed specifications of the operations are taken as given. In fact, for many purposes, the opera-

---

Presented at the 1984 Conference on Decision and Control, Las Vegas, NV, December 12–14, 1984. Stanley B. Gershwin and Sanjoy K. Mitter are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139. Richard R. Hildebrant is with the Charles Stark Draper Laboratory, Cambridge, MA 02139. Rajan Suri was with the Division of Applied Science, Harvard University, when this work was done. He is currently at the Department of Industrial Engineering, University of Wisconsin-Madison, Madison, WI 53706.

0272-1708/86/0400-0003 $01.00 © 1986 IEEE

tions themselves may be treated as black boxes.

The issue here is to move parts to machines in a way that reduces unnecessary idle time of both parts and machines. The loading problem is choosing the times at which the parts are loaded into the system or subsystem. The routing problem is to choose the sequence of machines the parts visit, and the scheduling problem is to choose the times at which the parts visit the machines.

The important considerations in routing include the set of machines available that can do the required tasks. It is often not desirable to use a flexible machine to do a job that can be done by a dedicated one, since the flexible machine may be able to do jobs for which there are no dedicated machines.

In scheduling, one must guarantee that parts visit their required machines while also guaranteeing that production requirements are met. At this level, the issue is allocating system resources in an efficient way. These resources include machines, transportation elements, and storage space.

A control problem at this level is to limit the effect of disruptions on factory operations. Disruptions are due to machine failures, operator absences, material unavailability, surges in demand, or other effects that may not be specified in advance but which are inevitable. This problem may be viewed as analogous to the problem of making an airplane robust to sudden wind gusts, or even to loss of power in one of three or more engines.

Factory Level

At each higher level, the time scale lengthens and the area under concern grows. At the next higher level, one must treat several cells. For example, in printed circuit fabrication, the first stage is a set of operations that prepares the boards. Metal is removed, and holes are drilled. At the next stage, components are inserted. The next stage is the soldering operation. Later, the boards are tested and reworked if necessary. Still later, they are assembled into the product. This process takes much time and a good deal of floor space.

Issues of routing and scheduling remain important here. However, setup times become crucial. That is, after a machine or cell completes work on one set of parts of the same or a smaller number of types, it is often necessary to change the system configuration in some way. For example, one may have to change the cutters in a machine tool. In printed circuit assembly, one must remove the remaining components from the insertion machines and replace them with a new set for the next set of part types to be made. The scheduling problem is now one of choosing the times at which these major setups must take place. This is often called the *tooling problem*.

Other issues are important at still longer time scales. One is to integrate new production demands with production already scheduled in a way that does not disrupt the system. Another class of decisions is that pertaining to medium-term capacity, such as the number of shifts to operate and the number of contract employees to hire for the next few months. Another decision, at a still longer time scale, is the expansion of the capital equipment of the factory. At this time scale, one must consider such *strategic goals* as market share, sales, product quality, and responsiveness to customers.

Discipline

Specified operating rules are required for complex systems. Manufacturing, communication, transportation, and other large systems degenerate into chaos when these rules are disregarded or when the rules are inadequate. In the manufacturing context, all participants must be bound by the operating discipline. This includes the shop-floor workers, who must perform tasks when required, and managers, who must not demand more than the system can produce. It is essential that constraints on allowable control actions be imposed on all levels of the hierarchy. These constraints must allow sufficient freedom for the decision makers at each level so that choices that are good for the system as a whole can be made, but they must not be allowed to disrupt its orderly operation.

Capacity

An important element in the discipline of a system is its capacity. Demands must be within capacity or excessive queuing will occur, leading to excessive costs and, possibly, to reduced effective capacity. High-level managers must not be allowed to make requirements that exceed their capacity on their subordinates; subordinates must be obliged to accurately report their capacities to those higher up.

All operations at machines take a finite amount of time. This implies that the rate at which parts can be introduced into the system is limited. Otherwise, parts would be introduced into the system faster than they could be processed. These parts would then be stored in buffers (or worse, in the transportation system) while waiting for the machines to become available, resulting in undesirably large work in process and reduced effective capacity. The effect is that throughput (parts actually produced) may drop with increasing loading rate when loading rate is beyond capacity. Thus, defining the capacity of the system carefully is a very important first step for on-line scheduling.

An additional complication is that manufacturing systems involve people. It is harder to *measure* human capacity than machine capacity, particularly when the work has creative aspects. Human capacity may be harder to *define* as well, since it can depend on circumstances such as whether the environment is undergoing rapid changes.

Defining, measuring, and respecting capacity are important at *all* levels of the hierarchy. No system can produce outside its capacity, and it is futile, at best, and damaging, at worst, to try. On the other hand, it may be possible to expand the capacity of a given system by a learning process. This is a goal of the Japanese *just-in-time* (JIT) approach, which takes place over a relatively long time scale.

It is essential, therefore, to determine what capacity is, then to develop a discipline for staying within it, and finally to expand it.

Uncertainty

All real systems are subject to random disturbances. The precise time or extent of such disturbances may not be known, but some statistical measures are often available. For a system to function properly, some means must be found to desensitize it to these phenomena.

Control theorists often distinguish between random events and unknown parameters, and different methods have been developed to treat them. In a manufacturing system, machine failures, operator absences, material shortages, and changing demands are examples of random events. Machine reliabilities are examples of parameters that are often unknown. Desensitization to uncertainties is one of the functions of the operating discipline. In particular, the system's capacity must be computed while taking disturbances into account, and the discipline must restrict requirements to within that capacity. The kinds of disturbances that must be treated differ at different levels of the time-scale hierarchy: at the shortest time scale, a machine failure influences which part is loaded next; at the longest scale, economic trends and technological changes influence marketing decisions and capital investments.

It is our belief that such disturbances can have a major effect on the operation of a plant. Scheduling and planning must take these events into account, in spite of the evident difficulty in doing so.

Feedback

In order to make good decisions under uncertainty, it is necessary to know something about the current state of the system and to use this information effectively. At the shortest time scale, this includes the conditions of the machines and the amount of material already processed. Control engineers know that designing good feedback strategies is generally a hard problem. It is essential, especially at the short time scale, that these decisions are calculated quickly and be relevant to long-term goals. The trade-off between optimality and computation gives rise to many interesting research directions.

SURVEY AND CRITIQUE OF PRACTICAL METHODS

The manufacturing environment is one of the richest sources of important and challenging control problems of which we are aware. Until recently, however, the classical and modern control community has not been attracted to this opportunity. One reason is undoubtedly that the manufacturing area has never been perceived as needing the help. Extreme competition from overseas manufacturers has, more than anything else, changed this perception.

Another reason the manufacturing area has not enjoyed the attention of control theorists is that the area has not been, and some argue is still not, amenable to their techniques. This is because, in part, modeling large complex systems is difficult. Also, there has not been sufficient information available for feedback control that is current or even correct. Control theory has, to a large extent, implicitly assumed a plant that is automatically controlled; manufacturing systems are run largely on manual effort. All this is beginning to change, however, due to the availability of inexpensive computation, the installation of more fully automated systems, and the additional requirements of flexibility, quality, etc., that are placed on these systems.

A wide variety of methods that deal with scheduling and planning are available to industry. The purpose of this section is to survey these methods and to critique them according to the outline of the previous section. A representative survey of current practice in controlling manufacturing systems is provided in this section. The intention is to give the reader perspective on the current state of manufacturing control.

Factory-Level Control

Traditional Framework

The manufacturing community is accustomed to thinking about production control within a particular mature framework. All the functions necessary for planning and executing manufacturing activities in order to make products most efficient have been grouped into a few large areas. These areas and the general interrelationships among them are shown in Fig. 1.

This diagram shows a tremendous amount of interaction, where information is fed forward and back, among the different areas. Also, the diagram deals mainly with the resource allocation aspect of the production control problem. Other important traditional areas that are integral to a successful control system are receiving, cost planning and control, and such financial functions as accounts receivable, accounts payable, etc.

Function Descriptions Brief descriptions of the functions performed within the major areas are given below:

Forecasting: Demand is projected over time horizons of various lengths. Different forecast models are maintained by this function.

Master Scheduling provides "rough-cut" capacity requirements analysis in order to determine the impact of production plans on

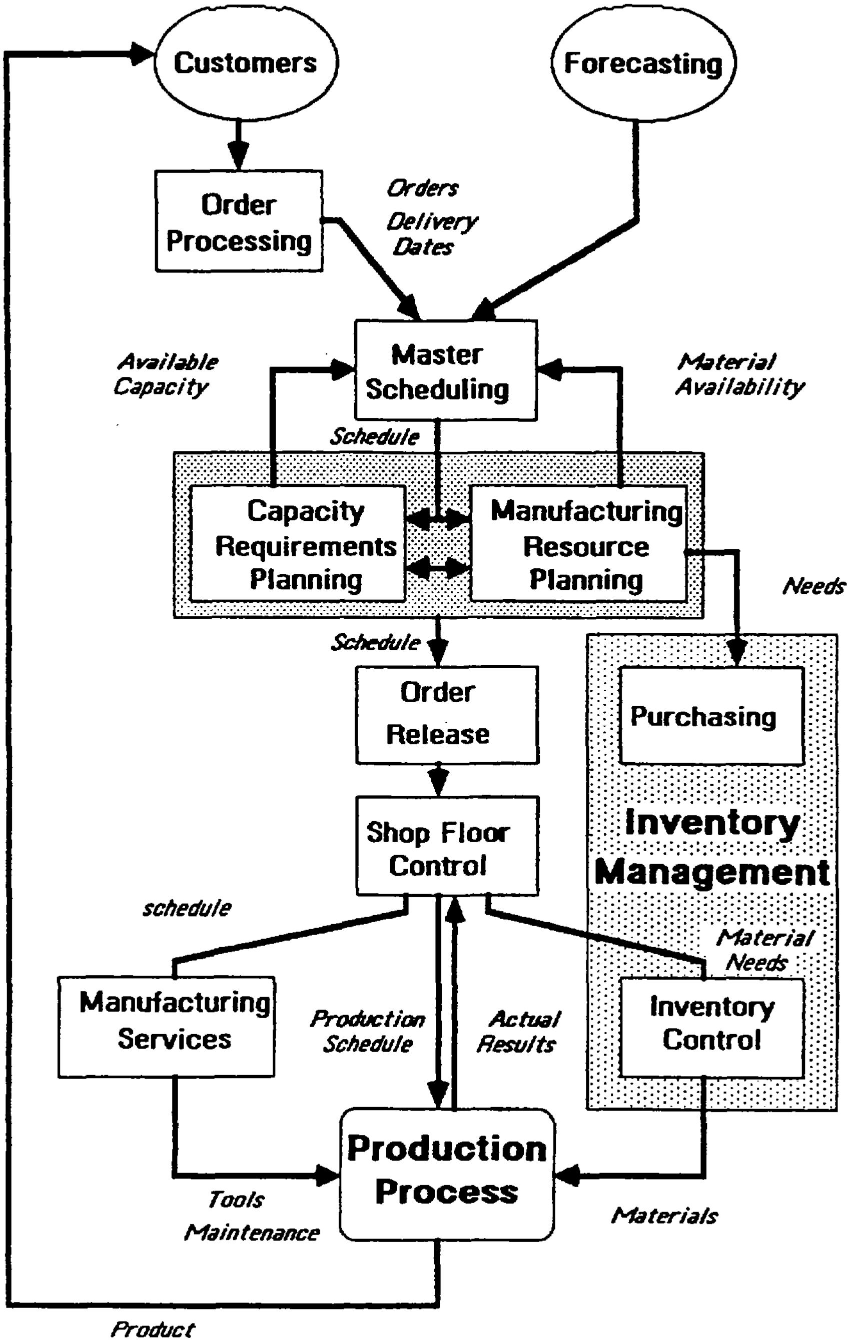


Fig. 1. Traditional framework in production control.

plant capacity. Comparisons are made between the forecast and actual sales order rate, sales orders and production, and, finally, scheduled and actual production.

Material Requirements Planning (MRP) determines quantity and timing of each item required — both manufactured and purchased. For each end-item, the quantity of all components and subassemblies is determined, and, by working backward from the date of final assembly, MRP determines when production or ordering of these subassemblies should occur. A more detailed capacity requirements analysis is made, and operation sequencing is determined. Also, lot-sizing is performed at this stage. While MRP tends to be highly detailed, there is no mechanism to take random events and unknowns into account; so it can lead to excess computer usage, misleading precision, delays in providing schedules, and rigidity.

Capacity Requirements Planning forecasts workcenter load for both released and planned orders and compares this figure with available capacity. The user specifies the number and duration of time periods over which the analysis is performed.

Order Release is the connection between manufacturing planning and execution. When an order is scheduled for release, this function creates the documentation required for initiating production.

Shop Floor Control is a lower level, "real-time" control function that is responsible for carrying out the production plan. This function performs priority dispatching and tracking of the product, as well as ancillary material and tooling. Data are collected on the disposition of the product and the performance of workcenters (utilization, efficiency, and productivity).

Inventory Control performs general accounting and valuation functions, as well as controlling the storage location of materials. It also often supports priority allocation of material to products or orders, and aids in filling order requisitions.

Planning Procedure These functions have always been performed. Many companies are organized according to these areas. For example, separate dedicated groups of people are often given the responsibility of controlling inventory, planning master schedules, etc.

The advent of the computer age brought software products that mirror almost exactly the functional framework outlined above. It is possible to buy software that addresses each functional area. In fact, the software is usually modularized so that the system may be acquired piecemeal.

The difficulty of the planning process is illustrated with the experiences of a large manufacturer of bed linens. They fabricate a basic set of products: sheets, pillow cases, and accessories. However, because bed linens have almost become high-fashion items, they come in a wide variety of sizes, prints, and styles. The force that drives the manufacturing process is initially the long-term forecast, but what is important is the ability to satisfy customer demand in the short to medium term. Customer demand manifests itself as an order for a particular set of products to be delivered at a particular time.

The basic problem they have is that orders are not being satisfied even though the overall level of finished inventory is extremely high. The product is being made, but not the right product at the right time!

Considering the large number of different products produced, the fact that production resources must be shared among those products, the limits on production capacity, and the difficulty of modeling manufacturing system dynamics, it is not surprising that the bed linen manufacturer has problems. Many companies do. The manufacturers that are most successful at controlling their operations generally make a small number of products and are able to forecast customer demand fairly accurately.

Recent Trends in Production Control

Finite-Capacity Material Requirements Planning Traditional MRP has offered little more than a computerized method of keeping voluminous records on material, and the resulting resource, requirements. There has never been an attempt, in any but the most superficial way, to account for the actual resource capacity in production planning and control. It has always been handled in an iterative, ad hoc, manual fashion. The manual approach is often a frustrating and impossible task.

There is growing interest in devising better factory-level models that integrate actual resource capacity with production requirements. In fact, one or two products that claim this capability have come onto the market within the past few years.

Products that attempt to perform finite-capacity planning often meet mixed reviews because their treatment cannot be comprehensive. A model formulation and its associated optimization procedure can be specified in a relatively straightforward way, but solving the problem with finite computational resources is impossible. Practical approaches must reduce the problem by making, what often turn out to be, limited assumptions. Also, factory dynamics for different industries, while often similar, can be quite different. Consider the domestic manufacturer of candies and confection that became enamored over the dazzling performance reports of a particular finite-capacity scheduler. This scheduler can be quite adept at modeling and controlling factories where discrete parts are manufactured, but the candy manufacturer's process was continuous! No amount of hammering could bend the finite-capacity scheduler into a shape that would solve their problems.

The Just-in-Time or Kanban Approach The just-in-time (JIT), or kanban (Kb), approach to manufacturing control is a Japanese refinement to the approach discussed above. The objective of this recent trend in material control is to reduce the need for large, expensive inventories of materials and subassemblies. By requiring that external and internal suppliers deliver just the right items, at just the right place, at just the right time, this objective may be met.

Kanban is a particular control implementation for forcing a just-in-time philosophy. A *kanban* is a job ticket that accompanies a part through the assembly process. When the part is actually installed in an assembly or subassembly, the kanban is sent back to its source to trigger the production of a new part. The control variable is the number of kanban tickets in the system.

The high risks of interrupted production due to low inventories are somewhat mitigated by imposing a great degree of discipline on all facets of manufacturing. Maintenance procedures and scheduling must be tightened up, lest the flow of parts that are needed downstream stop. Outside suppliers must ensure high quality in order to reduce the need for elaborate, and inventory-producing, inspections. Also, very good predictability of transportation times and strong communication ties are required of suppliers that participate in a JIT program. The long-term benefits of this discipline can lead to productivity increases beyond the simple reduction of inventory carrying costs [63], [32].

Implementing the JIT philosophy usually results in smaller and more frequent deliveries of materials. This can exacerbate the still necessary task of inventory management. Although zero inventory is an appealing goal, it should be moderated to the extent that costs required to achieve it increases.

The JIT philosophy for production control is most applicable where production requirements are known and fixed far in advance, and where buffering is not required to smooth the unavoidable effects of process time variations. This last point is illustrated by the material flow associated with flexible manufacturing systems, job shops, or any other system where a variety of parts with wide variations in process times share the same resources. Even without machine failures, buffers are required to reap the maximum production.

The JIT approach works best in applications such as the assembly process for products with predictable sales (refrigerators, automobiles, etc.). The uncertainties in these applications are not high enough to require intermediate buffering in order to achieve the maximum production rate possible.

Cell-Level Control

Traditional Approach

There have been very few successful approaches to scheduling the activities in a cell. Simulation is one that is widely used to determine scheduling strategies, floor layout, and other planning problems. However, it is expensive in both human and computer time since simulations, to be credible, tend to be complex and require a great deal of data. Many simulation runs are required to make a decision; the decision parameter must be "tuned" until optimal, or at least satisfactory, behavior is found.

Recent Developments in Cell Design and Control

Recent developments in automation and new constraints on the "flexibility" of the manufacturing process are beginning to alter the traditional concept of a cell and how that cell is to be controlled. One direction of development, called *group-technology cells, flowlines,* or *cellular manufacturing,* was stimulated by reports of Japanese successes. A family of products with very similar operation sequences is manufactured from start to finish in a single cell. This is intended to lead to a simplification of product flow and scheduling, tighter coupling of operations, less inventory, and greater worker coordination.

A second, stimulated by advances in automation and control technology, is the *flexible manufacturing system,* which is described below. A good overview of cellular manufacturing concepts can be found in Black [6] or Schonberger [64].

Flexible Manufacturing System Control A modern example of a cell is a flexible manufacturing system (FMS), which consists of several machines and associated storage elements, connected by an automated materials handling system. It is controlled by a computer or a network of computers. The purpose of the flexibility and versatility of the configuration is to meet production targets for a variety of part types in the face of disruptions, such as demand variations and machine failures.

In an FMS, individual part processing is practical because of two factors: the automated transportation system and the setup or changeover time (the time required to change a machine from doing one operation to doing another), which is small in comparison with operation times. The combination of these features enables the FMS to rapidly redistribute its capacity among different parts. Thus, a properly scheduled FMS can cope effectively with a variety of dynamically changing situations.

The size of these systems ranges from approximately 5 to more than 25 machines. They are also specifically designed for the concurrent processing of a number of different parts (5 to 10 unique parts types is not unusual), each of which may require a variety of processing (milling, drilling, boring, etc.). An FMS of average size, built for a large manufacturer of agricultural equipment, is shown in Fig. 2. There are, altogether, 16 machines that are serviced by automatically guided vehicles (AGVs). An area for loading and unloading parts on and off the AGVs is set aside to one end of the system. Parts enter the system here, proceed through their respective process plan, and exit.

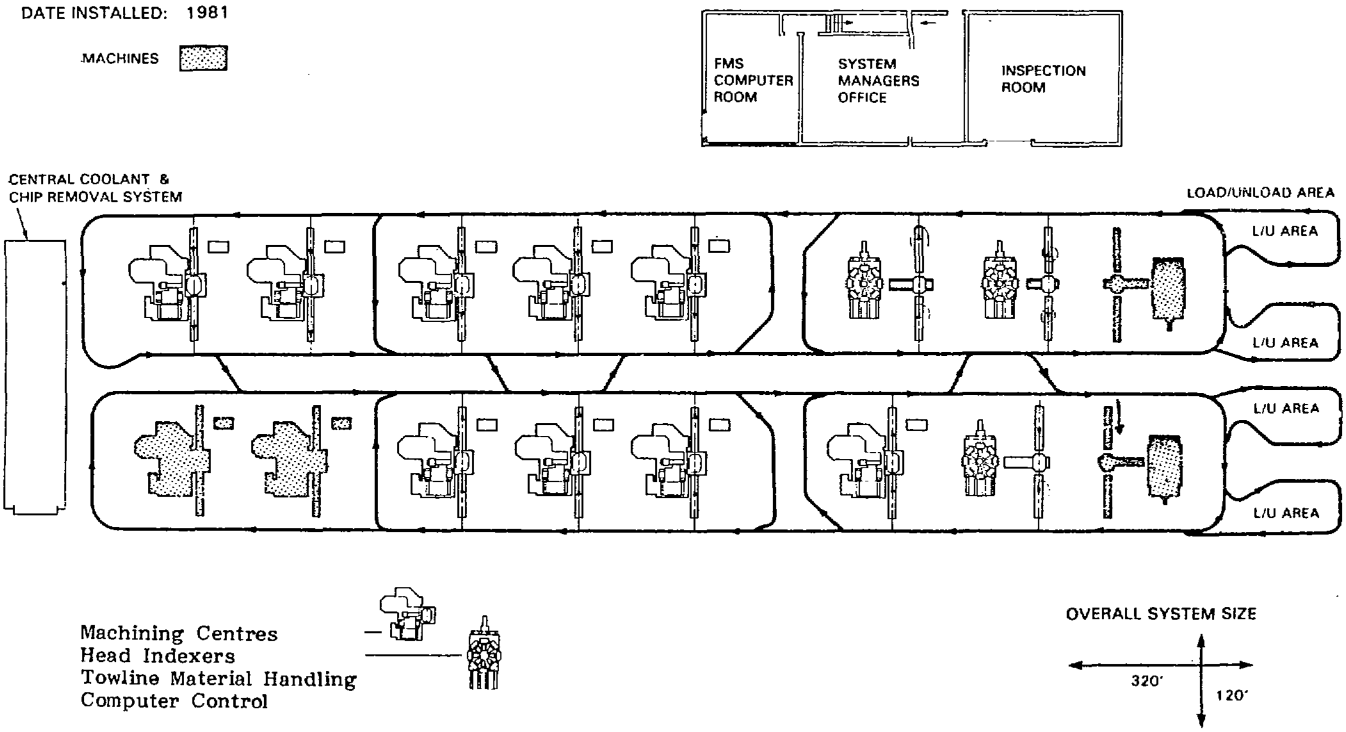


Fig. 2. Flexible manufacturing system of average size.

A flexible manufacturing system is a simple cell whose main objective is to meet a predefined master production schedule. The operational decisions that must be made include:

- Allocation of operations (and tools) to machines, such that the following, often conflicting, subobjectives must be met:
    - — Workload requirements are evenly balanced among the machines and material handling systems.
    - — Machine failures have a minimum effect on other machines' work availability.
    - — Work-in-process requirements are minimized.
    - — Processing redundancy (duplicate tooling) is maximized.
- Reallocation of operations and tools to machines when machines fail, such that, in addition to those objectives listed above, tool-changing effort is minimized.
- Real-time allocation of resources for processing pieceparts, such that:
    - — Workload requirements are evenly balanced among the machines and material handling systems.
    - — Quality of processed parts is maximized.

The first two areas are generally not well addressed by the vendor community and, in fact, often cause the users major operational problems when trying to run an FMS. Because of the difficulty in juggling the conflicting objectives under sometimes severe constraints (limits on the number of tool pockets per machine and on the weight the tool chain may bear), it is very difficult to manually allocate processing to resources. Some recent strides have been taken in solving this problem, but the capability is not yet widespread and has not yet been integrated into the operating software that controls FMSs.

The real-time scheduling of parts to machines, however, is addressed directly by the vendors that supply "turn-key" FMSs. Each vendor usually takes a unqiue approach to the scheduling problem (this is motivated, in part, by the unique aspects of each vendor's design) and, because of a perceived proprietary edge, is often reluctant to divulge the details of its implementation. Nevertheless, after analyzing the behavior of many FMSs over a period of time, we can make the following observations.

The decision classes, or control variables, for scheduling the activity in an FMS are listed below. In principle, one can construct a detailed schedule before the fact. In practice, however, the complexity of the problem (the large number of possible decision choices and uncertainty in material and resource availability) prevents this.

The general approach taken by the practitioners of FMS control is that of *dispatch scheduling*. Decisions are made as they are needed. Very little information is considered when making these decisions. The criteria and constraints for a variety of questions related to dispatch scheduling are:

- *Part sequence into FMS:* Since an FMS can process a number of different parts and since these parts are required in certain ratios relative to one another, active control of the part input sequence is required.
- *Sequencing of fixturings:* Many parts must make a number of passes through the system in order to process different sides. The sequence for these separate passes could be chosen to enhance the performance of the system.
- *Sequencing of operations:* Once in the system, a part must often visit a number of different machines before processing is complete. The sequence of these separate machine visits could be chosen to enhance the performance of the system.
- *Machine choice:* Often a particular operation may be performed at more than one machine. When this is true, a choice must be made among the possibilities.
- *Cart choice:* Many FMSs employ a number of separate carts for transporting parts from machine to machine. When the need arises for transporting a part, a choice among the carts of the system must be made.
    - — *cart movement:* Carts are always moving, except while undergoing load/unload operation or while queuing at an occupied node. Shortest routes are chosen when there is a destination. Deadlocks are checked for periodically.
    - — *requests for carts:* Intervals are computed and parts are introduced. Backlogs of parts are tracked. The closest cart with the correct pallet is chosen for loading parts. The closest empty pallet is chosen for each part coming off a shuttle.
- *Operation and frequency selection for quality check:* Many FMSs being built are equipped with a coordinate measuring machine (CMM). The purpose of this machine is to monitor the quality of the parts being processed as well as the processes themselves. Through the measurement of part dimensions, the nature of process errors (tool wear, machine misalignment, fixture misalignment, etc.) may be inferred. Because the CMM resource is limited, the intelligent selection of operations to measure and the frequency with which to measure them is required in order to ensure that quality standards are satisfied and that processing errors are quickly identified.

Machine-Level Control

The bottom tier of the manufacturing structure is comprised of individual workstations, which may be actual machinery or even lone workers (as is the case with manual assembly systems). Control at the machine level does not really include material flow, scheduling, or other logistical considerations. These issues have been accounted for at the cell and factory levels.

The problems encountered at this level are sometimes more in line with those that have been traditionally treated within the classical and modern control framework. The domain is often continuous, rather than discrete, and there is often opportunity for instrumenting the machinery for full automatic control and feedback.

The Traditional Approach

In the beginning, there were just hand tools. All control and feedback was accomplished through eye-hand coordination. This continued to be the case, for the most part, up until recently (1950s). The tools (lathes, drill presses, etc.) became larger and more complex, but the principle remained the same. Then computers were applied and numerically controlled (NC) machinery was the result. Here the position, feed, and speed of the tool relative to the part is controlled through standard feedback techniques. In addition, the different operations a part required could be programmed to occur automatically on one machine in the proper sequence.

Operation sequencing is generally performed open loop: there has not been sufficient reason to alter the sequence. This is changing in some environments where there is full automation. It may happen that a tool breaks part way through a "tape segment." If the part has to leave the machine and come back for any reason (quality-control check, extract broken tool, etc.), it is difficult to pick up where the processing left off.

Recent Developments

Until recently, the position of the tool, and its feed rate and speed relative to the part, has been controlled in an entirely open-loop manner. Regardless of what was happening (wearing of tools, anomalies in casting dimensions and quality, etc.), these variables would remain constant. This is beginning to change. By monitoring the power requirements of a particular cut, the condition of the casting/tooling combination can be determined and adjustments made.

Electronic vision is another means by which feedback is being used in control at the machine level. These systems check for the presence or absence of tools in the spindle. Other techniques for measuring the wear on these tools are also being employed.

SURVEY AND CRITIQUE OF RECENT RESEARCH

This section reviews recent developments in analysis and optimization of manufacturing systems. It is intended for control engineers who wish to become familiar with research in this field. We should emphasize at the outset that there is a large body of literature available on traditional approaches to manufacturing. For example, the area of production and operations management (POM) occupies a significant place in most business schools, and many textbooks exist for this well-developed area [85]. Here we will restrict ourselves to the *systems* aspects of manufacturing problems, to areas relevant to the framework as discussed earlier, and to recent developments in these areas, which we believe could significantly affect the progress of the field.

We use the time-scale hierarchy mentioned, rather than the framework of practical methods. The former is more appealing from the control point of view, and perhaps more amenable to rigorous development.

We adopt the distinction, as proposed in Suri [78], between *generative* and *evaluative* techniques or models. A generative technique is one that takes a set of criteria and constraints, and generates a set of decisions. An evaluative technique is one that takes a set of decisions and predicts (evaluates) the performance of a system under those decisions. (The terms *prescriptive* or *normative,* and *descriptive,* are also used for these two categories.)

While we concentrate on recent developments, it is appropriate to comment briefly on early research in manufacturing systems that can be found in the management science and operations research literature. Much of this was directed at production planning and scheduling problems.

In particular, a great deal of the work on generative techniques for production scheduling and planning was concerned with the mathematical problem of fitting together the production requirements of a large number of discrete, distinct parts [24]. Such combinatorial optimization problems are very difficult in the sense that they often require an impractical amount of computer time. Furthermore, they are limited to deterministic problems so that random effects, including machine failures and demand uncertainties, cannot be analyzed. An excellent review of production scheduling methods, including the use of heuristics and hierarchical approaches to solve large problems, can be found in Graves [29].

The early work on evaluative models was mainly an attempt to represent the random nature of the production process by using queuing-theoretic models, such as the classic M/M/1 and M/G/1 queues [52].

The applicability of queuing theory to manufacturing was considerably enhanced by the development of network-of-queues theory [46], [28]. However, only the more recent development of efficient computational algorithms and good approximation methods has enabled the implementation of reasonable "first-cut" evaluative models of fairly complex manufacturing systems, as described later.

Another early development in the area of evaluative models was the use of computer-based simulation methods, which employ a "Monte Carlo" approach to system evaluation. With the growing accessibility of computing power, the development of easy-to-use simulation packages, and the advancement of simulation theory, this area has made major strides forward recently.

Long-Term Decisions

In this section, we consider decisions that involve considerable investment in plant, equipment, or new manufacturing methods. Typically, such decisions may take over a year to implement and may have an operational lifetime of 5 to 20 years during which they are expected to pay back.

Generative Techniques

Traditional systems-based approaches for generating long-term decisions include the production planning and hierarchical approaches mentioned above, as well as strategic planning, forecasting, decision analysis, and location analysis. We do not deal with these here, but an overview and literature survey can be found in [77].

In the context of automated manufacturing, mathematical programming techniques (LP and IP) have been applied to selection of equipment and of production strategies [30], [73], [87]. However, the constraints involved in the mathematical programming problem formulations can be very complex. Whitney [86] has proposed *sequential decisions,* which is a new framework for developing heuristic algorithms for solving these complex optimization problems. It has been successfully applied to the problem of selecting parts and equipment for manufacturing in a very large organization.

Some recent approaches, which should be of interest to the control community, use dynamic investment models for long-term decision making. The decision to invest in alternative manufacturing strategies (and equipment), over a period of time, is formulated as an optimal control problem [7], [25], [53]. Such models offer qualitative insight to help decision makers faced with the complex set of investment alternatives that modern manufacturing systems involve. However, practical application and use of these models remains to be seen.

Evaluative Techniques

The evaluation of long-term effects of a decision on an enterprise is a particularly difficult problem, and evaluative models for long-term planning deal primarily with strategic and accounting issues [45], [47], [53]. Strategic issues involve such questions as how improved product quality or response time to orders will affect the market share. Accounting issues require models to trade off current expenditures with future (uncertain) revenue streams. Neither of these areas is of primary interest to the current audience. However, we should mention two factors. The first is that evaluative models for strategic and accounting issues are currently undergoing radical changes, in the face of the (relative) failure of U.S. industry to make prudent investments [1], [57].

The second important point, which is often missed by those undertaking modeling/analysis studies, is that the long-term decisions are influenced to a large extent by these strategic and accounting issues. Even though we do not cover them here, it is important for analysts working in this general area not to lose sight of the forest for the trees. Many modeling and analysis efforts fail to be useful to the manufacturing community because they focus on minor technical points and do not provide the overall insight that is needed for this stage of the planning process. Professor Milton Smith (of Texas Technological University) said at the recent First ORSA/TIMS Conference on Flexible Manufacturing Systems that around a hundred man-years had been expended on solving minimum makespan scheduling problems, but he did not know of a single company that used minimum makespan to schedule their shop-floor operations!

Medium-Term Decisions

Here we are concerned with a time period ranging anywhere from a day to a year, and the scope of the decisions involves primarily trade-offs between different modes of operation, but with only minor investments in new equipment/resources.

Traditionally, such decisions have been the domain of master scheduling, MRP, and inventory management systems. These systems generally do not account for uncertainty in a direct manner, but rather, in an indirect way through the use of "safety" values, whether in stocks, lead times, or other quantities. Master scheduling and MRP systems work to a deterministic plan, which gets updated periodically (say once a week or once a month); see [51] and [84] for insights into these points.

Inventory theory models and analyzes the effects of uncertainty to derive optimal stock policies. Inventory stocking policies assume that each item stocked has an exogenous demand, modeled by some stochastic process, and attempt to find the best stocking policy for each item.

The fact that the demand on inventory comes as a result of the master scheduling and MRP decisions is ignored, and thus it is clear that much useful information for decision making is being thrown away. Of course, it is the size and complexity of a manufacturing system that makes it very difficult to solve the entire problem simultaneously. Nevertheless, we feel that suitable structures can be developed to make the decision making more coherent across these components. Some attempts in this direction are described in this section.

Generative Techniques

We begin by reviewing generative techniques for this level of decision making. Control theorists are familiar with the concept of time-scale decomposition and hierarchical control, and should therefore readily understand the idea behind hierarchical production planning. It partitions the problem into a hierarchy of subproblems, with successively shorter time scales. The solution of each subproblem imposes constraints on lower subproblems. The advantages of the hierarchical approach are many: in addition to computational savings, this approach requires less detailed data, and it mimics the actual organizational structure [58].

The original ideas for this approach based the hierarchical structure on intuitive and heuristic arguments. However, control theorists should find it interesting that Graves [30] showed, by the use of Lagrange multipliers, that the Hax-Meal hierarchy could be derived as a natural decomposition of a primal optimization problem. An alternative hierarchy, based not on optimality but rather feasibility considerations, is derived by Suri [74], also using multiplier methods.

These hierarchical approaches assume that demand and capacity are known and deterministic over a period of time, and then re-solve the planning problem periodically. Recent developments in manufacturing systems have sought to represent uncertainty explicitly in the problem formulation. This uncertainty includes not just demand but also equipment failures (hence, randomly varying capacity). Since this usually leads to an intractable problem, the contribution of the new approaches is primarily in ways that they propose to formulate the problem or approximate the solution.

Hildebrant and Suri [36] proposed a hierarchical procedure where the hierarchy is derived from heuristic arguments based on tractability considerations, but the interaction between the levels is based on a mathematical programming problem. To get around the difficulty of solving a stochastic optimization problem, they proposed an "open-loop feedback" policy where the dynamics of the system between failure states is replaced by a static average of the time spent in each failure state (or each capacity condition). The technique showed reasonable improvement over existing heuristics [35].

Kimemia [49] and Kimemia and Gershwin [50] have derived an alternative, closed-loop solution to this problem. Their approach has also been to separate the relatively long-term issues (the response to machine failures and to production backlogs and surpluses) from the short-term problem of part dispatching. The long-term problem is modeled as a continuous dynamic programming problem. A feedback control law, which determines the next part to be loaded and when it should be loaded as a function of current machine state and current production surplus, is sought.

This formulation, which reflects the disruptive nature of machine failures, had previously been proposed by Olsder and Suri [61], but they had concluded that it was too hard to solve exactly. The contribution of Kimemia and Gershwin has been to find a good approximation to the exact solution. Essentially, this involves two steps in a dynamic programming framework: separating the top-level problem (the solution to a Bellman equation) into a number of subproblems, obtained formally through a constraint-relaxation procedure, and then approximating the value function for each subproblem by a quadratic.

More recent work by Gershwin, Akella, and Choong [27] has further simplified the computational effort. Simulation results indicate that the behavior of a manufacturing system is highly insensitive to errors in the cost-to-go function, so the Bellman equation can be replaced by a far simpler procedure.

Evaluative Models

Evaluative models for this decision-making level involve both analytic approaches and simulation. Important features are the ability to represent production uncertainties (such as machine failures) and limited buffer stocks, in order to trade off between the two. For large systems, this is, again, analytically intractable. The earliest work in this field is surveyed by Koenigsberg (1959). Notable contributions were made by Buzacott [8]–[10], who looked at various approximate analytic models that gave insight into these issues.

Most of these analytic studies are based on the Markov models of transfer lines and other production systems. An appreciation for the difficulty of the problem is seen from the fact that the largest general model for which an exact solution is available involves three machines with two buffers between them [26].

A promising, recent development has come out of a technique for decomposing a production line into a set of two-machine one-buffer subsystems [26]. This idea had been previously proposed by Zimmern [90] and Sevastyanov [68], but an efficient and accurate method had not been developed. Gershwin's procedure for solving this system is analogous to the idea of solving two-point boundary value problems. Numerical results indicate that the method is very accurate, and, what is more, fairly large problems (20 machines) can be solved in reasonable time. The technique is, however, currently restricted to the case where the cycle times of the machines are deterministic and equal. Altiok [3] has recently developed methods for systems with more general phase-type processing time distributions.

Other evaluative techniques include queuing network models, and simulation. Both of these methods can be used for short-term decision making as well. However, we feel that queuing network models are best suited to more aggregated decision making, while simulation is more suited to detailed decisions. Therefore, we discuss the former here and the latter in the next subsection, although the particular application may suggest the use of one or the other technique for either of these levels.

A fairly recent development in (analytic) evaluative models of manufacturing systems has been the growing use of *queuing network* models for system planning and operation. A simple-minded, static, capacity allocation model does not take into account the system dynamics, interactions, and uncertainties inherent in manufacturing systems. Queuing network models are able to incorporate some of these features, albeit with some restrictions, and thus enable more refined evaluation of decisions for manufacturing systems. The increased use of such models stems primarily from the advances made in the computational algorithms available to solve queuing networks, both exactly and approximately.

Buzen's algorithm [14] made the solution of these systems tractable. Solberg [71] applied this to capacity planning for FMS, and Stecke [72] used it for solving production planning problems. Shanthikumar [70] developed a number of approximate queuing models for manufacturing. The development of the *mean value analysis* (MVA) technique for solving these networks (Reiser and Lavenberg) opened up a host of new extensions and approximations. Various approximate MVA algorithms have been developed [66], [4], which enable fast and accurate solution of very large networks. Hildebrant and Suri [36] and Hildebrant [35] applied MVA techniques to both design and real-time operation problems in FMS.

An extension [81] enabled efficient solution of systems with machine groups. Another recent extension, called *priority mean value analysis* (PMVA) [69], allows a wide variety of operational features to be modeled.

An important reason for the increasing popularity of such models in manufacturing is that they have proved their usefulness in the area of computer/communication systems modeling, in terms of giving reasonable performance predictions. Recent analysis has given a basis to the robustness of queuing network models for use in practical situations [76].

The disadvantages of queuing network models are that they model many aspects of the system in an aggregate way, and they fail to represent certain other features, such as limited buffer space. (Some recent developments, e.g., Buzacott and Yao [13] and Suri and Diehl [80] do allow limited buffer sizes.) The output measures they produce are average values, based on a steady-state operation of the system. Thus, they are not good for modeling transient effects due to infrequent but severe disruptions such as machine failures.

However, the models tend to give reasonable estimates of system performance, and they are very efficient: that is, they require relatively little input data, and do not use much computer time. A typical FMS model [81] might require 20 to 40 items of data to be input, and run in 1 to 10 sec on a microcomputer, in contrast to the much larger numbers for simulation. Thus, these models can be used interactively to quickly arrive at preliminary decisions. More detailed models can then refine these decisions.

Queuing network models suggest themselves for use in the middle level of a hierarchy. Development of queuing network models along with suitable control aspects to tie in the lower and higher levels of the hierarchy could be a useful topic of research.

Lasserre [55] and Lasserre and Roubellat [56] represent the medium-term production planning problem as a linear program of special structure, and develop an efficient solution technique for it.

Short-Term Decisions

Decisions at this level typically have a time frame of from a few minutes up to about a month. Traditional generative models have included those for lot sizing and scheduling, using both exact approaches as well as heuristics or rules. There have been a number of recent, interesting developments in this area, which are now described.

Traditional lot-sizing models traded off the cost of setting up a machine with the cost of holding inventory, on an individual product basis. The Japanese (just-in-time and kanban) approaches have challenged these concepts as being narrow-minded and myopic in terms of the long-term goals of the organization. They advocate operation with minimal or no inventory, claiming that this not only saves inventory carrying costs but also gives rise to a learning process that leads to more balanced production in the long run [63], [32].

This thesis is becoming more widely accepted in U.S. industries as well. However, reduced inventory leads to line stoppages and inefficiencies in the short term. It is, therefore, logical to ask what is the optimal rate of reducing inventory so that short-term losses are traded off against long-term gains due to increased learning. There has been some preliminary investigation of this point [79]. It is a problem that would fit naturally into an optimal control framework, and further investigation would be useful.

There have also been some recent studies indicating that lot sizing in a multi-item environment ought to be treated as a vector optimization problem. The idea is that the lot size of each product affects the production rate of other products, primarily through the queuing of each lot of parts waiting for other lots to be done at each machine. Therefore, one ought to consider the joint problem of simultaneously optimizing all the lot sizes. This integer programming problem would normally be computationally intractable for any realistic manufacturing system. However, by modeling the system as a queuing network and then solving a resulting nonlinear program, some recent results have been obtained [48]. This is a promising development that needs further exploration.

Hitz [37], [38] studied the detailed, deterministic scheduling of a special class of flexible manufacturing systems: flexible flow shops. In these systems, parts follow a common path from machine to machine. He found that by grouping parts appropriately, he could design a periodic sequence of loading times. This substantially reduced the combinatorial optimization problem.

Erschler, Roubellat, and Thomas [21] describe a deterministic, combinatorial scheduling technique that searches for a *class* of optimal decisions. Rather than deciding which part to send into the system next, it presents to the user a set of candidate choices. This flexibility is intended as a response to the random events such as machine failures that are difficult to represent explicitly in a scheduling model.

Perhaps the most widely used evaluative tool for manufacturing systems today is *simulation* [65]. The term "simulation" in this context refers specifically to computer-based discrete event simulation. Such a model mimics the detailed operation of the manufacturing system, through a computer program that effectively steps through each event that would occur in the system (or to be more precise, each event that we wish to model).

In principle, simulation models can be made very accurate—the price is the programming time to create the model, the input time to generate detailed data sets, and the computer time each time the model is run. In addition, the more phenomena that the analyst tries to represent, the more complex the code, and the more likely there are errors, some of which may never be found.

It is sometimes forgotten that the accuracy of a simulation is limited by the judgment and skill of the programmer. Detail and complexity are not necessarily synonymous with accuracy, if major classes of phenomena are left out. (While simulation can be used at any of the levels of the decision-making process, we choose to describe it here since it can examine the most detailed operation of a manufacturing system.)

Two reasons for the recent popularity of simulation are the number of software tools that have been developed to make simulation more accessible to manufacturing designers, and the decrease in computing costs and the availability of microcomputers. These factors make it well worth an organization's effort to use simulation before making large investments. In addition, there have been many developments in the design and analysis of simulation experiments, which have contributed to the acceptance of simulation as a valid and scientific methodology in this field.

Recent developments in software tools for simulation can be categorized into simulation languages, "canned packages," interactive model development (or graphical input), and animation (or output graphics). The two input/output graphics features will not be discussed further here, but good examples are SIMAN (for input) and SEE WHY (for output).

Although simulation languages have been around for a while, the last five years have seen the development of many powerful languages, such as GPSS/H, SIMSCRIPT II.5, and SLAM II, as well as the development of languages specially tailored to the manufacturing user (e.g., SIMAN and MAP/1). Also, most languages are now available on microcomputers as well (e.g., GPSS/PC, SIMSCRIPT II.5, SIMAN, MICRONET).

Another development, specially geared to the manufacturing designer, has been the development of canned packages, which do not require programming skills, but are completely data driven (e.g., GCMS, GFMS, SPEED). Of course, they have a number of structural assumptions built into them, in terms of how the manufacturing system operates, but can be useful for very quick analysis of a system. At the other end of the spectrum, for very detailed simulation, it may be necessary to resort to a programming language such as FORTRAN or PASCAL. See [5] for a discussion of the trade-offs among these options.

It should be noted that simulation is useful for planning, as well as off-line analysis of operating strategies. However, its structure and computation-time requirements make it currently unsuitable for on-line decision making. Even though simulation is perhaps the most widely used computer-based performance evaluation tool for manufacturing systems, we would recommend greater use of analytic and queuing network models prior to conducting the more expensive simulation studies — in comparison to the numbers quoted for queuing network models, a simulation model might require 100 to 1000 data items and 15 to 10,000 sec to run on a microcomputer.

In the area of simulation design and analysis, there have been several developments that should be of interest to the control community. The analysis of simulation outputs — which involves parameter identification, confidence interval generation, detection of bias and initial transients, and run length control — has used many techniques from time-series analysis and spectral methods (see [54] for a survey). Parameter optimization in simulations involves stochastic approximation techniques [40], [59], [83], which again are familiar ground to our community.

A recent development, called *perturbation analysis of discrete event systems*, enables very efficient optimization of parameters in simulations (see [44] for a survey). This technique is related to linearization of dynamic systems, and, again, has parallels with conventional dynamic systems [39], [41]. Essentially, it enables the gradient vector of system output with respect to a number of parameters to be estimated by observing only one sample path. In this sense, it is an evaluative and "semigenerative" tool, since it not only evaluates decisions but also suggests directions for improving the decisions.

While much of the original work on perturbation analysis relied on experimental results to demonstrate its accuracy [42], [43], recent analyses have given it a more rigorous foundation [75] and also proved that it is probabilistically correct for certain systems [17], [40], [83], as well as better than repeated simulation [16], [89].

Another recent, interesting development has been the application of the Petri net theory to the performance analysis of manufacturing systems [23]. In the past, the main use of Petri nets (in computer science) was to answer such qualitative questions as: Will there be any deadlocks? However, there have been some important recent advances in the theory of timed Petri nets.

Following some work by Cunningham-Greene [20], Cohen et al. [18], [19] have developed a linear systems theoretic view of production processes. This enables efficient answers to some complex performance questions. It also gives rise to a parallel set of control-theoretical concepts for discrete events systems, e.g., transfer functions, controllability, observability. The main disadvantages it has currently are that it can only deal with completely deterministic situations and that it is only evaluative, not generative. However, it is a promising new development.

One of the most useful areas requiring more research is that of real-time control of manufacturing systems, at a detailed level. Little theoretical research has been done on this, apart from the large body of heuristics that exists for scheduling [29]. A few researchers have treated the issues in a formal way [11], [2], [15]. This seems to be an area for control theorists to apply their expertise.

Indeed, Ho et al. [44] have coined the term DEDS, for discrete event *dynamic* system, to emphasize that manufacturing systems are a class of dynamic systems, and that there are concepts from dynamic system theory that need to be developed or applied for DEDS as well. In the past, we have seen DEDS analyzed either by purely probabilistic approaches (e.g., Markov chains, queues) or by purely deterministic approaches (scheduling and other combinatorial methods). The work by Cohen et al., as well as the perturbation analysis approach, have shown the use of a dynamic systems view of the world.

As an example, Suri and Zazanis [89] have used perturbation analysis combined with stochastic approximation to adaptively optimize a queuing system. This could be used, for example, for improving the choice of lot sizes for a number of different parts, while a facility is operating — the approach is simple to implement and has obvious applications in real systems. However, many interesting questions of convergence, etc., remain to be answered for this adaptive method.

CONCLUSION

We have described a framework for many of the important problems in manufacturing systems that need the attention of people trained in control and systems theory. We have shown how existing practical methods solve those problems, and where they fall short. We have also shown how recent and on-going research fits into that framework. An important goal of this effort has been to encourage control theorists to make the modeling and analysis efforts that will lead to substantial progress in this very important field.

References

[1] R. A. Abbott and E. A. Ring, "The MAPI Method — Its Effects on Productivity: An Alternative Is Needed," vol. 2, no. 1, pp. 15–30, 1983.

[2] R. Akella, Y. Choong, and S. B. Gershwin, "Performance of Hierarchical Production Scheduling Policy," *IEEE Trans. on Compon., Hybr., and Manuf. Tech.*, Sept. 1984.

[3] T. Altiok, "Approximate Analysis of Exponential Tandem Queues with Blocking," *Eur. J. Oper. Res.*, vol. 11, pp. 390–398, 1982.

[4] Y. Bard, "Some Extensions to Multiclass Queuing Network Analysis," in *Performance of Computer Systems* (M. Arato, ed.), North Holland, Amsterdam, 1979.

[5] J. P. Bevans, "First, Choose an FMS Simulator," *American Machinist*, pp. 143–145, May 1982.

[6] J. T. Black, "An Overview of Cellular Manufacturing Systems and Comparison to Conventional Systems," *Industrial Engineering*, vol. 15, no. 11, pp. 36–40, Nov. 1983.

[7] M. C. Burstein and M. Talbi, "Economic Justification for the Introduction of Flexible Manufacturing Technology: Traditional Procedures Versus a Dynamics Based Approach," *Annals of Oper. Res.*, 1985.

[8] J. A. Buzacott, "Automatic Transfer Lines with Buffer Stocks," *Int. J. Prod. Res.*, vol. 5, no. 3, pp. 183–200, 1967.

[9] J. A. Buzacott, "The Role of Inventory Banks in Flow-Line Production Systems," *Int. J. Prod. Res.*, vol. 9, no. 4, pp. 425–436, 1971.

[10] J. A. Buzacott, "The Production Capacity of Job Shops with Limited Storage Space," *Int. J. Prod. Res.*, vol. 14, no. 5, pp. 597–605, 1976.

[11] J. A. Buzacott, "Optimal Operating Rules for Automated Manufacturing Systems," *IEEE Trans. on Auto. Contr.*, vol. AC-27, no. 1, pp. 80–86, Feb. 1982.

[12] J. A. Buzacott and J. G. Shanthikumar, "Models for Understanding Flexible Manufacturing Systems," *AIIE Trans.*, pp. 339–350, Dec. 1980.

[13] J. A. Buzacott and D. D. W. Yao, "Flexible Manufacturing Systems: A Review of Models," Working Paper No. 7, Dept. of IE, Univ. of Toronto, Mar. 1982.

[14] J. P. Buzen, "Computational Algorithms for Closed-Queuing Networks with Exponential Servers," *C. ACM*, vol. 16, no. 9, pp. 527–531, Sept. 1973.

[15] C. Cassandras, "A Hierarchical Routing Control Scheme for Material Handling Systems," *Annals of Oper. Res.*, 1985.

[16] X. R. Cao, "Convergence of Parameter Sensitivity Estimates in a Stochastic Experiment," *Proc. IEEE Conf. Dec. and Contr.*, Dec. 1984.

[17] X. R. Cao, "Sample Function Analysis of Queuing Networks," *Annals of Oper. Res.*, 1985.

[18] G. Cohen, D. Dubois, J. P. Quadrat, and M. Viot, "A Linear System-Theoretic View of Discrete-Event Processes," *Proc. 22nd IEEE Conf. Dec. and Contr.*, Dec. 1983.

[19] G. Cohen, D. Moller, J. P. Quadrat, and M. Viot, "Linear System Theory for Discrete Event Systems," *Proc. 23rd IEEE Conf. Dec. and Contr.*, Dec. 1984.

[20] R. A. Cunningham-Greene, "Describing Industrial Processes and Approximating Their Steady-State Behaviour," *Op. Res. Q*, vol. 13, pp. 95–100, 1962.

[21] J. Erschler, F. Roubellat, and V. Thomas, "Aide à la Décision dans L'Ordancement d'Atelier en Temps Réel," 3è me Journées Scientifiques et Techniques Automatisée, ADEPA, Toulouse, June 1981.

[22] Draper Lab., *The Flexible Manufacturing Systems Handbook*, The Charles Stark Draper Laboratory, Cambridge, Massachusetts, 1982.

[23] D. Dubois and K. E. Stecke, "Using Petri Nets to Represent Production Processes," *Proc. 22nd IEEE Conf. Dec. and Contr.*, Dec. 1983.

[24] B. P. Dzielinski and R. E. Gomory, "Optimal Programming of Lot Sizes, Inventory, and Labor Allocations," *Manag. Sci.*, vol. 11, no. 9, pp. 874–890, July 1965.

[25] C. Gaimon, "The Dynamic, Optimal Mix of Labor and Automation," *Annals of Oper. Res.*, 1985.

[26] S. B. Gershwin, "An Efficient Decomposition Method for the Approximate Evaluation of Production Lines with Finite Storage Space," MIT Laboratory for Information and Decision Systems, Rep. LIDS-R-1309, Dec. 1983.

[27] S. B. Gershwin, R. Akella, and Y. C. Choong, "Short-Term Production Scheduling of an Automated Manufacturing Facility," MIT Laboratory for Information and Decision Systems, Rep. LIDS-FR-1356, Feb. 1984.

[28] W. J. Gordon and G. F. Newell, "Closed-Queuing Systems with Exponential Servers," *Op. Res.*, vol. 15, no. 2, pp. 254–265, Apr. 1967.

[29] S. C. Graves, "A Review of Production Scheduling," *Op. Res.*, vol. 29, no. 4, pp. 646–675, 1981.

[30] S. C. Graves, "Using Lagrangian Techniques to Solve Hierarchical Production Planning Problems," *Manag. Sci.*, vol. 28, no. 3, pp. 260–275, Mar. 1982.

[31] S. C. Graves and B. W. Lamar, "A Mathematical Programming Procedure for Manufacturing System Design and Evaluation," *Proc. IEEE Int. Conf. on Cir. and Compu.*, 1980.

[32] R. W. Hall, *Zero Inventories*, Dow-Jones Irwin, 1983.

[33] A. C. Hax and H. C. Meal, "Hierarchical Integration of Production Planning and Scheduling," in *Logistics* (M. A. Geisler, ed.), North Holland, 1975.

[34] P. Heidelberger and P. D. Welch, "A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations," *C. ACM*, vol. 24, no. 4, pp. 233–245, Apr. 1981.

[35] R. R. Hildebrant, "Scheduling Flexible Machining Systems When Machines Are Prone to Failure," Ph.D. Thesis, MIT, 1980.

[36] R. R. Hildebrant and R. Suri, "Methodology and Multilevel Algorithm Structure for Scheduling and Real-Time Control of Flexible Manufacturing Systems," *Proc. 3rd Int. Sym. on Large Engin. Sys.*, Memorial Univ. of Newfoundland, pp. 239–244, July 1980.

[37] K. L. Hitz, "Scheduling of Flexible Flow Shops," MIT Laboratory for Information and Decision Systems, Rep. LIDS-R-879, Jan. 1979.

[38] K. L. Hitz, "Scheduling of Flexible Flow Shops — II," MIT Laboratory for Information and Decision Systems, Rep. LIDS-R-1049, Oct. 1980.

[39] Y. C. Ho, "A Survey of the Perturbation Analysis of Discrete Event Dynamic Systems," *Annals of Oper. Res.*, 1985.

[40] Y. C. Ho and X. R. Cao, "Perturbation Analysis and Optimization of Queuing Networks," *JOTA*, vol. 40, no. 4, pp. 559–582, 1983.

[41] Y. C. Ho and C. Cassandras, "Computing Costate Variables for Discrete Event Systems," *Proc. 19th IEEE Conf. Dec. and Contr.*, Dec. 1980.

[42] Y. C. Ho, M. A. Eyler, and T. T. Chien, "A New Approach to Determine Parameter Sensitivities of Transfer Lines," *Manag. Sci.*, vol. 29, no. 6, pp. 700–714, 1983.

[43] Y. C. Ho, M. A. Eyler, and T. T. Chien, "A Gradient Technique for General Buffer Storage Design in a Serial Production Line," *Int. J. Prod. Res.*, vol. 17, no. 6, pp. 557–580, 1979.

[44] Y. C. Ho, R. Suri, X. R. Cao, G. W. Diehl, J. W. Dille, and M. A. Zazanis, "Optimization of Large Multiclass (Nonproduct Form) Queuing Networks Using Perturbation Analysis," *Large-Scale Systems*, 1984.

[45] G. K. Hutchinson and J. R. Holland, "The Economic Value of Flexible Automation," *J. of Manuf. Sys.*, vol. 1, no. 2, pp. 215–228, 1982.

[46] J. R. Jackson, "Jobshoplike Queuing Systems," *Manag. Sci.*, vol. 10, no. 1, pp. 131–142, Oct. 1963.

[47] R. Jaikumar, "Flexible Manufacturing Systems: A Managerial Perspective," Harvard Business School, Working Paper, 1984.

[48] U. S. Karmarkar, S. Kekre, and S. Kekre, "Lotsizing in Multi-Item Multi-Machine Job Shops," Univ. of Rochester, Grad. School of Mgt., Working Paper QM8402, Mar. 1984.

[49] J. G. Kimemia, "Hierarchical Control of Production in Flexible Manufacturing Systems," MIT Laboratory for Information and Decision Systems, Rep. LIDS-TH-1215.

[50] J. G. Kimemia and S. B. Gershwin, "An Algorithm for the Computer Control of Production in Flexible Manufacturing Systems," *IEE Trans.*, vol. 15, no. 4, pp. 353–362, Dec. 1983.

[51] O. Kimura and H. Terada, "Design and Analysis of Pull System: A Method of Multistage Production Control," *Int. J. Prod.*

Res., vol. 19, no. 3, pp. 241-253, 1981.

[52] L. Kleinrock, *Queuing Systems*, John Wiley, 1975.

[53] N. Kulatilaka, "A Managerial Decision Support System to Evaluate Investments in FMSs," *Annals of Oper. Res.*, 1985.

[54] A. M. Law, "Statistical Analysis of Simulation Output Data," *Oper. Res.*, vol. 31, no. 6, pp. 983–1029, Dec. 1983.

[55] J. B. Lasserre, "Étude de la Planification à Moyen Terme d'une Unite de Fabrication," thesis presented to Université Paul Sabatier for the degree of Docteur Ingenieur, 1978.

[56] J. B. Lasserre and F. Roubellat, "Une Methode Rapide de Resolution de Certaines Programmes Lineaires a Structure en Escalier," *RAIRO Oper. Res.*, vol. 14, no. 2, pp. 171–191, May 1980.

[57] L. C. Leung and J. M. A. Tanchoco, "Replacement Decision Based on Productivity Analysis — An Alternative to the MAPI Method," *J. Manuf. Sys.*, vol. 2, no. 2, pp. 175–188, 1983.

[58] H. C. Meal, "Putting Production Decisions Where They Belong," *Harvard Business R.*, vol. 62, no. 2, pp. 102–111, Mar. 1984.

[59] M. S. Meketon, "Optimization in Simulation: A Tutorial," presented at Winter Simulation Conf., WSC83, Dec. 1983.

[60] G. J. Olsder and R. Suri, "Time-Optimal Control of Flexible Manufacturing Systems with Failure Prone Machines," *Proc. 19th IEEE Conf. Dec. and Contr.*, Dec. 1980.

[61] *Operations Research*, Special Issue on Simulation, vol. 31, no. 6, 1983.

[62] M. Reiser and S. S. Lavenberg, "Mean Value Analysis of Closed Multichain Networks," *J. ACM*, vol. 27, pp. 313–323, Apr. 1980.

[63] R. J. Schonberger, *Japanese Manufacturing Techniques*, Free Press, 1982.

[64] R. J. Schonberger, "Integration of Cellular Manufacturing and Just-In-Time Production," *Industrial Engineering*, vol. 15, no. 11, pp. 66–71, Nov. 1983.

[65] T. Schriber, "The Use of GPSS/H in Modeling a Typical Flexible Manufacturing System," *Annals of Oper. Res.*, 1985.

[66] P. J. Schweitzer, "Approximate Analysis of Multiclass Closed Networks of Queues," presented at Int. Conf. on Stochastic Contr. and Optimization, Amsterdam, 1979.

[67] A. Seidman and P. J. Schweitzer, *Real-Time On-Line Control of a FMS Cell*, Working Paper QM8217, Grad. School of Manag., Univ. of Rochester, New York, 1982.

[68] B. A. Sevastyanov, "Influence of Storage Bin Capacity on the Average Standstill of a Production Line," *Theory Prob. Appl.*, pp. 429–438, 1962.

[69] S. Shalev-Oren, A. Seidman, and P. J. Schweitzer, "Analysis of Flexible Manufacturing Systems with Priority Scheduling: PMVA," *Annals of Oper. Res.*, 1985.

[70] J. G. Shanthikumar, "Approximate Queuing Models of Dynamic Job Shops," Ph.D. Thesis, Dept. of IE, Univ. of Toronto, 1979.

[71] J. J. Solberg, *CAN-Q User's Guide*, Report No. 9 (Revised), School of IE, Purdue Univ., W. Lafayette, Indiana, 1980.

[72] K. E. Stecke, *Production Planning Problems for Flexible Manufacturing Systems*, Ph.D. Dissertation, School of IE, Purdue Univ., W. Lafayette, Indiana, 1981.

[73] K. E. Stecke, "Formulation and Solution of Nonlinear Integer Production Problems for Flexible Manufacturing Systems," *Manag. Sci.*, vol. 29, no. 3, pp. 273–288, Mar. 1983.

[74] R. Suri, *Resource Management Concepts for Large Systems*, Pergamon Press, 1981.

[75] R. Suri, "Robustness of Queuing Network Formulae," *J. ACM*, vol. 30, no. 3, pp. 564–594, July 1983a.

[76] R. Suri, "Infinitesimal Perturbation Analysis of Discrete Event Dynamic Systems: A General Theory," *Proc. 22nd IEEE Conf. Dec. and Contr.*, Dec. 1983b.

[77] R. Suri, "Quantitative Techniques for Robotic Systems Analysis," Chapter in *Handbook of Industrial Robotics* (S. Y. Nof, ed.), Wiley, 1984b.

[78] R. Suri, "An Overview of Evaluative Models for Flexible Manufacturing Systems," *Annals of Oper. Res.*, 1985.

[79] R. Suri and S. DeTreville, "Getting from Just-In-Case to Just-In-Time: Insights from a Simple Model," *J. Oper. Manag.*, 1985.

[80] R. Suri and G. W. Diehl, "A Variable Buffer Size Model and Its Use in Analyzing Closed Queuing Networks with Blocking," *Manag. Sci.*, 1985.

[81] R. Suri and R. R. Hildebrant, "Modeling Flexible Manufacturing Systems Using Mean Value Analysis," *J. Manuf. Sys.*, vol. 3, no. 1, pp. 27–38, 1984.

[82] R. Suri and C. K. Whitney, "Decision Support Requirements in Flexible Manufacturing," *J. Manuf. Sys.*, vol. 3, no. 1, pp. 61–69, 1984.

[83] R. Suri and M. A. Zazanis, "Perturbation Analysis Gives Strongly Consistent Estimates for the M/G/1 Queue," *Manag. Sci.*, to appear in 1986.

[84] T. Tabe, R. Muramatsu, and Y. Tanaka, "Analysis of Production Ordering Quantities and Inventory Variations in a Multistage Ordering System," *Int. J. of Prod. Res.*, vol. 18, no. 2, pp. 245–257, 1980.

[85] R. J. Tersine, *Production and Operations Management*, North Holland, 1980.

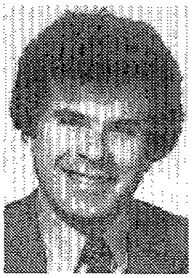
[86] C. K. Whitney, "Control Principles in Flexible Manufacturing," submitted to *J. Manuf. Sys.*, 1984.

[87] C. K. Whitney and R. Suri, "Algorithms for Part and Machine Selection in Flexible Manufacturing Systems," *Annals of Oper. Res.*, 1985.

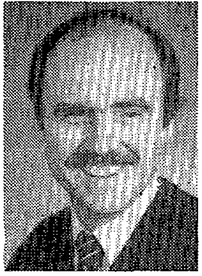
[88] D. D. W. Yao and J. A. Buzacott, "Modeling a Class of Flexible Manufacturing Systems with Reversible Routing," Tech. Rep. 8302, Dept. of IE, Univ. of Toronto, Aug. 1983.

[89] M. A. Zazanis and R. Suri, "Comparison of Perturbation Analysis with Conventional Sensitivity Estimates for Regenerative Stochastic Systems," submitted to *Oper. Res.*, 1985.

[90] B. Zimmern, "Etudes de la Planification des arrets aleatoires dans les chaines de production," *Revue de statistique appliquee*, vol. 4, pp. 85–104, 1956.

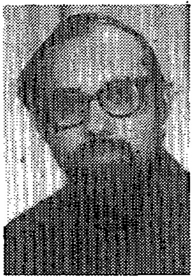


Stanley B. Gershwin received the B.S. degree in engineering mathematics from Columbia University, New York, in 1966, and the M.A. and Ph.D. degrees in applied mathematics from Harvard University, Cambridge, Massachusetts, in 1967 and 1971. He is a Principal Research Scientist at the MIT Laboratory for Information and Decision Systems and a Lecturer in the MIT Department of Electrical Engineering and Computer Science. He is also Assistant Director of the MIT Laboratory for Information and Decision Systems, and President of Technical Support Software, Inc. (TSSI). In 1970–1971, he was employed by the Bell Telephone Laboratories in Holmdel, New Jersey, where he studied telephone hardware capacity estimation. At the Charles Stark Draper Laboratory in Cambridge, Massachusetts, from 1971 to 1975, Dr. Gershwin investigated problems in manufacturing and in transportation. His interest in these areas, as well as control, optimization and estimation, continues at MIT and at TSSI. He has studied traffic assignment, the measurement of traffic-flow parameters, and related areas in transportation. He has investigated routing optimization in networks of machine tools and the effect of limited buffer storage space on transfer lines and assembly networks. During June 1981, Dr. Gershwin was the guest of the French scientific agency (CNRS) and visited laboratories in Toulouse, Bordeaux, Paris, and Valenciennes to observe their research in manufacturing systems.



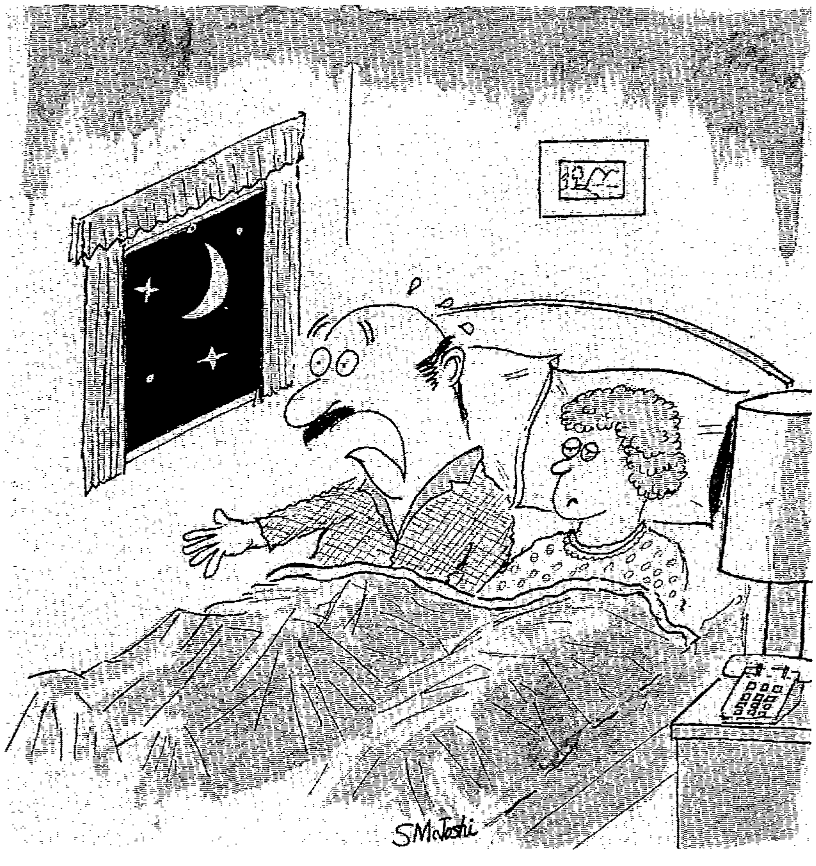
Richard R. Hildebrant is a member of the research staff at the Charles Stark Draper Laboratory in Cambridge, Massachusetts, where he is Chief of the Industrial Operations Methods Section. His background in modern and classical control theory has been applied to the design and analysis of aircraft, navigation, and other aerospace systems. More recently, his professional interests have focused on the application of operations research, AI, computer science, manufacturing engineering, and related disciplines to the development of scheduling and control software for manufacturing systems. His expertise, ranging from the design, analysis, and specification of complex manufacturing and assembly systems to the development and implementation of real-time scheduling and control software, has been employed by numerous major corporations undergoing both large and small CAM projects. Dr. Hildebrant received his Ph.D. from the Massachusetts Institute of Technology in 1980.

Rajan Suri is currently Associate Professor of Industrial Engineering at the University of Wisconsin-Madison. He received his bachelor's degree (1974) from Cambridge University (England) and his M.S. (1975) and Ph.D. (1978) from Harvard University. In 1981, he received the IEEE Donald P. Eckman Award for outstanding contributions in his field. His current interests are in modeling and decision support for manufacturing systems, specializing in flexible manufacturing systems. He is the author of many journal publications and several books, and has chaired international conferences on this subject. He is a consultant to major industrial corporations and also a principal of Network Dynamics, Inc., a Cambridge Massachusetts, firm specializing in software for modeling manufacturing systems.



Sanjoy Mitter was born in Calcutta, India, and educated in India and England. He received a Ph.D. in automatic control from Imperial College of Science and Technology, University of London, in 1965. He has taught at Case Western Reserve University and has been at MIT since 1969, where he is currently Professor of Electrical Engineering and Director, Laboratory for Information and Decision Systems. He has held visiting appointments at numerous institutions in Asia, Europe, and the USA, including the School of Mathematics, Tata Institute of Fundamental Research, Bombay, India; Scuola Normale Superiore, Pisa, Italy; Imperial College, London; and INRIA, France. He is a Fellow of the IEEE. Dr. Mitter has made extensive contributions in research in the broad area of systems and control, statistical signal processing, and mathematical problems related to the above areas.

Out of Control



"There I was, walking on the bridge between theory and practice, and suddenly, out of nowhere, comes this gigantic, strongly positive, self-adjoint operator with a dense domain, and smashes the bridge all to pieces! What a nightmare!"