<Survey Paper>

Journal of the Korean Institute of Industrial Engineers
Vol. 2 No. 1 June 1976

SCHEDULING THEORY AND PROBLEMS: REVIEW AND CATEGORIZATION OF SOLUTION PROCEDURES

Sung Hyun Park*

1. INTRODUCTION

The purpose of this paper is twofold. The first objective is to introduce the elements of the theory of scheduling, to classify the scheduling problems, and to review the recent developments in the area of scheduling. The second objective is to convey to the unfamiliar reader the basic philosophy behind the theory of scheduling and its tremendous applicability to various activities of everyday life as well as a wide variety of modern sciences.

Suppose we have a set of well defined jobs (tasks, projects, items, and the like) and a certain amount of resources (processors, machines, facilities, and the like) to perform the jobs. Scheduling may be defined as the process of constructing an ordering (priority) of the jobs and allocating the resources over time to perform the ordered jobs. Since the early develop ments of scheduling were mainly motivated by problems arising in manufacturing, much of the research literature in scheduling uses the terminology of manufacturing: resources are usually called machines, places where the machines are ready to process the jobs are called shops, and a number of related and dependent activites that consist of a job are called operations.

The theory of scheduling is applicable not only in manufacturing areas, but in non-manufacturing areas such as transportation, communication, services, etc. Scheduling problems arise quite naturally in various phases of everyday life: for instance, patients waiting on test facilities in a hospital, programs to be run at a computing center, bank customers at a row of tellers' windows, repair jobs waiting at a service station, and many others. Scheduling problems seem to appear everywhere and overwhleming. Note that, in scheduling terminology, patients, programs, bank customers and repairs are considered as jobs, and test facilities, computing centers, tellers and service stations are considered as machines.

The vital elements in scheduling are machines and jobs. Machines are characterized in terms of their qualitative or quantitative availability, and jobs are described in terms of their machine requirements, durations (processing times), and the times at which they may be started and they are due. In some instance, there are technological restrictions on the order in which jobs can be performed. Therefore, solving a scheduling problem under a given objective criterion amounts to answering two kinds of questions:

1. In what order should the jobs be performed?

2. Which machines will be allocated to perform each job and when will each job be performed?

In other words, the essence of scheduling problems gives rise to two decisions: decisions on job sequencing (the order of jobs), and decisions

* College of Business and Industry
Missippi State University

on machine allocating. Notice that scheduling problems presume that the nature of the jobs to be scheduled has been completely described and the configuration of the machines available has been determined. In fact, determining the type of machines to be used and the number of jobs to be scheduled is not the question of scheduling, but the question of planning. Planning is the fundamental managerial decision-making which must precede the scheduling decision. Questions of scheduling, in a sense, are perhaps not as vital as decisions on planning. Nevertheless, if proper selection of schedule can yield some incremental improvement, it seems pointless to neglect the opportunity.

The typical assumptions that are usually made in scheduling problems are:

1. All assigned jobs must be performed (no job cancellation) and no job may be on two ma chines at the same time. That is, any two operations of a job cannot be executed at the same time.

2. All machines employed in performing the jobs must be entirely specified, and no machine may process more than one operation at a time.

3. The set of operations required to perform each job must be completely defined, including any restrictions on the order in which these must be performed. Also each operation, once started, must be performed to completion without interruption.

Althoughrelaxation of one or more assumptions above is possible, most of the research literature in scheduling seems to impose more assumptions and concentrate on developing algorithms for particular cases. To make the essence of scheduling problems clear, an example will be consid ered in the following section.

2. AN EXAMPLE

Suppose there are two jobs to be scheduled on three machines where each job consists of exactly 3 operations, one operation on each

		Machine 1	Machine 2	Machine 3
Job	1	4	1	2
	2	1	4	3

Table 1 : Processing Times

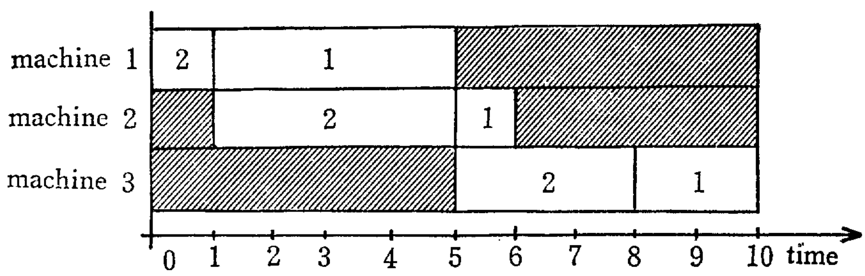

Figure 1 Schedule 1

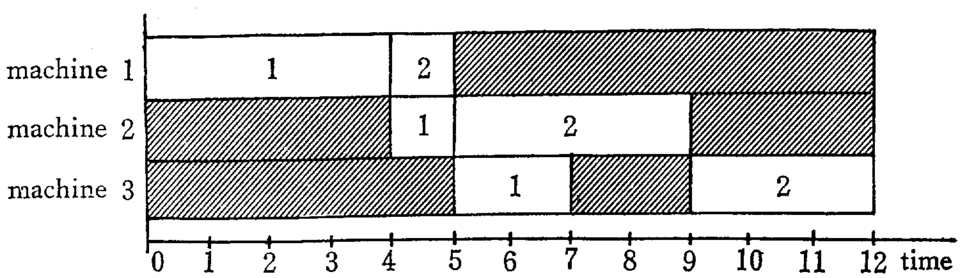

Figure 2 Schedule 2

machine. The processing time of each operation is given in Table 1. The objective criterion is to finish the jobs as early as possible. If we assu me that the machines are ready at time zero and we are willing to start each operation as early as possible, there are total 8 schedules to consider (i.e., two subsequences, jobs 1-2 or 2-1 on each machine, and, since there are 3 machines, we have total 2^3 sequences). Two out of 8 schedules are shown below.

The completion time for schedule 1 is $t=12$ and for schedule 2 it is $t=10$. Obviously, schedule 2 is preferred to schedule 1. In general, suppose there are n jobs and m machines, and each job must be processed on each machine exactly once. Since each machine must process n operations, the number of possible subsequences is $n!$ for each machine. If the subsequences on each ma chine were entirely independent, there would be total $(n!)^m$ sequences to consider. If $n=6$ and $m=3$, for such small number of jobs or machines the total number of sequences to consider is a surprizing number, 373, 248, 000. Clearly, complete enumeration of the total sequences is out of the question even for modern computers, and this is the agony of most scheduling problems. We have discussed a particular scheduling model, and as mentioned earlier, there are a wide variety of scheduling models. An attempt will be made to classify the models in the next section.

3. SCHEDULING MODELS

Until mid 1950's, there has been virtually no reported attempt to treat scheduling problem- from the point of view of analytical models with the explicit purpose of optimizing some measure of performance. In the beginning, there was great reluctance to use this family of valuable optimizing techniques because it was new. Soon, however, the study of scheduling theory has attracted researchers of a rather high calib er, and there has been atremendous growth in the use and application of scheduling theory. To classify the major scheduling models, it seems necessary to characterize the following five types of information:

1. The nature of the job arrivals.
2. The nature of the job parameters.
3. The number of machines.
4. The manner in which job assignments can be made.
5. The criterion by which a schedule will be evaluated.

The nature of the job arrivals provies the distinction between static and dynamic problems. In a static problem a certain number of jobs ar rive simultaneously in a shop that is available for work and no further jobs will arrive. In a dynamic problem jobs arrive intermittently and will continue to arrive into the future. Static models have proven more tractable than dynamic models and have been subjected to more extensive study by many researchers.

The nature of the job parameters provides the distinction between deterministic and stochastic (probabilistic) problems. The parameters of each job are its processing time, engineering content, technical importance, etc. An excellent exposition of dynamic and stochastic models can be found in the book by Conway et al. [10].

The third type of information describes the number of machines in the shop. In general, problems are divided into two: those with a single machine and those with multiple machines. The single machine case is a specialized scheduling model in which there is only a single resource and an ordering of the jobs completely determines a schedule. As elementary as it may be, however, its study is very important. The analysis of single machine problems has frequently uncovered valuable insights that are useful in more complicated multiple machine cases. It is therefore a building block in the development of a comprehensive uederstanding of scheduling concepts, an understanding that should ultimately facilitate the modeling of complicated systems. The list of recent references of single machine cases are lengthy; the following studies are cited as examples only: Baker and Martin [4], Gelders and Kleindorfer[17], McMahon and Florian [25], Moore [26], Park and Heiser [32], Rinnooy Kan et al. [33], Shwimer [36], and

Sidney [37]. The process of scheduling in multiple machines requires both machine allocation decisions and sequencing decisions (i.e., the ord er of jobs on the specified machine). As would be expected, multiple machine systems exhibit almost unlimited varieties of scheduling models. If the models are divided according to the manner in which job assignments can be made, we have three major models: parallel machines, flow shops and job shops.

The scheduling problem in parallel machines is that of one-operation-per-job scheduling on identical machines (i.e., parallelism in resource structure). That is, each job is to be processed once and only once by any one of the several identical machines. A review of literature of this case can be found in Park[30]. In the case of flow shop the jobs consist of operations in which a special precedence structure applies. Thus each job requires a specific sequence of operations to be carried out by machines. Ingeneral, in the flow shop there are m different machines, and each job consistsof m operations, each of which requires a different machine and all the jobs foll ow essentially the same path from one machine to another. Note that the example in Section 2 is a flow shop scheduling problem, if each job must follow machines 1, 2 and 3 in order. Some of the recent literature on this subject are: Ashour [1], Baker [2], Burns and Rooker[7], Dutta jand Cunningham [14], Gupta [18], Smith et al. [38], and Uskup and Smith [40].

The job shop is a general case of the flow shop. The difference is that in the job shop a job may require only a subset of machines to process its entire operations, and a job may req uire processing by the same machine more than once. Therefore, the important distinction of the job shop from the flow shop is that there is no common pattern of job movement from machine to machine. Although it is easy to state the general job shop problem, it is extremely difficult to solve this problem, and this is still one of unsolved problems in scheduling. Job shop problems have been recently discussed by Balas [5], Chanlton and Death [9], Florian et al. [16], and Hollloway and Nelson [20, 21], and others.

Figure 3 provides a bird's eye view of the classification of scheduling systems as of 1976. In Figure 3 the ditto for 'dynamic' means that dyna mic systems have two branches exactly like'static' and, similarly, the ditto of 'probabilistic' has the same two branches as 'deterministic'.

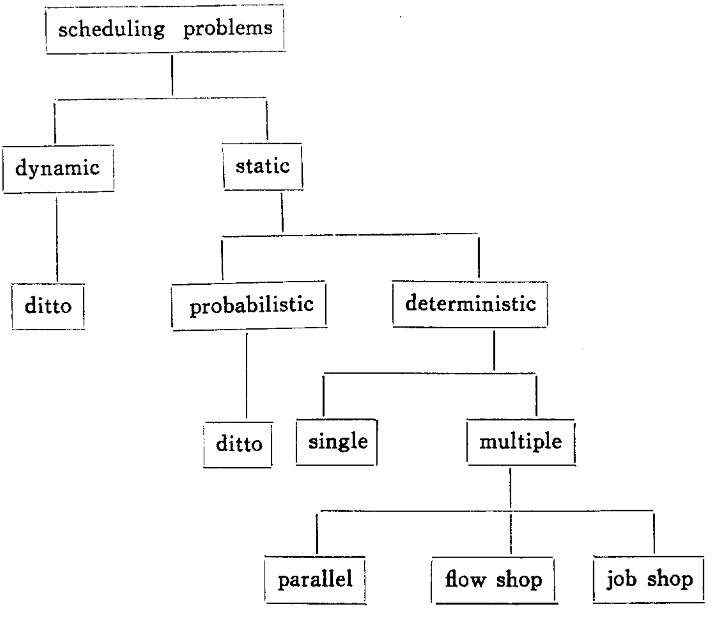

Figure 3 Bird's eye view of scheduling problems

The criteria are measures of schedule evaluation, and an optimal schedule cannot be advanced withot an approporiate criterion associated with a scheduling problem. Ideally, the criterion should consist of all costs that relate with scheduling decisions. In practice, such costs are often difficult to identify and even harder to measure. However, a study of the criteria proposed in the literature reveals that, in general, there are three principal types of costs that can be affected by scheduling decisions. These are the costs of

1. utilization of resources,
2. inventory,
3. job lateness.

The efficient utilization of resource requires scheduling decisions which will provide a given work load to be accomplished with a smaller aggregate demand on facilities. It produces a short schedule time and permits a given facility to do more work. The in-waiting inventory is another economic consequence of scheduling decisions, which depends on the position of individual jobs in the schedule. The effort of reducing inventory provides the incentive for minimizing average flow-time of jobs, and also provides rapid response to demands for competitive sales advantage. The cost of job lateness is obvious in many situations, especially competitive manufacturing with due-date. The proposed criteria may be summarized as follows. They are the minimization of:

mean of flow-times
mean of weighted flow-times
maximum makespan (length of time required to complete all jobs)
maximum tardiness (job lateness after due-date)
total cost of tardiness
weighted sum of tardiness
total cost of production
total setup time or cost
total cost of in-waiting inventory
number of tardy jobs
etc.

For any scheduling problem with a given criterion, very frequently more than one problem-solving-technique have been proposed to find an optimal schedule. Indeed, the scheduling field has been mainly concerned with the development, application andevaluation of problem-solving-techniques.

4. METHODOLOGIES

A survey of the techniques shows that there are basically six different methodologies:

1. general mathematical programming techniques
2. branch and bound methods
3. combinatorial procedures
4. network methods
5. heuristic solution procedures
6. simulation techniques.

Depending on the nature of the model and the compelxity of the problem, each method has its own advantages and disadvantages, and the selection of an appropriate technique often requires a great deal of study. If we utilize advantages of each technique, it is also conceivable that hybrid methods (i.e., combination of more than one technique) can solve some problems efficiently.

General mathematical programming techniques include linear, dynamic, integer, convex programming and Lagrangian methods. A merit of these techniques is that almost all scheduling problems may be formulated by a mathematical programming and can be solved. However, the computational effort required to solve the problem grows at an exponential rate with increasing problem size, and developing an efficient code remains an intriguing question. Particularly, in situations where alternate optima are desired, efficient coding usually requires more complicated tactics than if only one optimum is desired. Nevertheless, for many problems in which other efficient optimizing procedures have not been developed (e.g., minimization of maximum tardiness in parallel machines, minimization of weighted number of tardy jobs on a single machine), mathematical programming should be a choice. The references are abundant for mathematical programming techniques, and the following studies are revealed in recent publications. They are: Baker [3], Dutta and Cunningham [14], Elmaghraby [15], Karp and Held [23], Kolesar et al. [24] and Zaloom and Vatz [41] among many others.

A clever and promising approach that attracts vigorous study is branch and bound methods, sometimes also known as "implicit enumeration" methods. In essence, the approach consists of two principal concepts: one is that all feasible schedules must be accounted for by complete enumeration either explicitly or implicitly, and the other is that the least number of schedules should be accounted for explicitly from the virtue of dominance, bounding or feasibility considerations of jobs. There are two important processes in any branch and bound method: branching and bounding. Branching stems from the fact that in terms of a tree (i.e., a scheduling prob lem may be viewed as a connected graph having no loops, for which an optimal path is sought.), the procedure is continually concerned with choosing a branch of the tree to evaluate a path. Bounding is the process of bounding the value of the objective function by calculating a lower bound (or an upper bound) at each node of the tree in order to curtail dominated paths and select a branch. Branch and bound methods have been applied by Baker [2], McMahan and Florian [25], Panwalkar and Khan [29], Park and Elmaghraby [31], Rinnooy Kan et al. [33], Shwimer [36], Uskup and Smith [40], and others.

Combinatorial procedures have, in essence, the same principle as the branch and bound methods: searching for a best sequence while enumerating all sequences. However, branching and boundrng processes are not associated. Instead, combinatorial methods rely on developing certain rules concerning optimal ordering of jobs to curtail a large portion of enumerations. The works by Burns and Rooker [7], Gupta [18], Root [34], and Sidney [37] belong to combinatorial procedures.

The techniques of notwork analysis (CPM and PERT) have been widely used in a variety of areas: planning, research and development, construction, marketing, etc. Since many of the scheduling problems can be visualized as networks, and analysed with the use of network concepts, it is appropriate to consider the network methods as aproblem-solving-technique of scheduling problems. The network model is essentially a means of representing precedence relationships between jobs under the assumption that there are unlimited resources. Therefore, if we impose limitation on resources, the network model is a scheduling model and, accordingly, the techniques of network analysis are valuable tools for scheduling problems. The list of literature of network scheduling is lengthy; the following studies are cited as examples only: Bennington and McGinnis [6], Davis [12], Davis and Patterson [13], Schrage [35], and Sidney [37].

The computational effort to solve scheduling problems for an optimal schedule grows remarkably fast as problem size increases. Where computing capacility is scarce or expensive, it is quite reasonable to consider heuristic solution procedures, which are capable of obtaining good solutions very quickly, but which do not necessarily yield optimal solutions. This heuristic solution approach was the basis for the papers by Campbell et al. [8], Davis and Patterson [13], Palmer[28], and Tilquin [39].

Finally, simulation methods have been widely employed mainly for dynamic or probabilistic models. In dynamic models jobs arrive at the shop randomly over time from a probability distribution of job-arrivals, so that computer simulation studies have been naturally involved. For probabilistic models analytic approaches are often quite difficult, and the study of simulation methods is an alternative choice. The simulation study has also been suggested in deterministic scheduling problems where good analytic solutions are not available, and it shows some success. Some of the references for simulation studies are Conway [11], Hershauer and Ebert [19], Hottenstein [22], Moore and Wilson [27], and Park and Heiser [32].

5. CONCLUDING REMARKS

The theory and problems in scheduling have been classified and reviewed. The manufacturing terminologies such as job, shop and machine were explicitly used to describe scheduling problems. Nevertheless, the theory of scheduling should be equally applicable to non-manufacturing areas. To the author's knowledge, the theory of scheduling and its applications are bound to flourish in research and practice for at least three good reasons. First, almost all management problems of large scale resolve themselves, is the final stage, into problems of scheduling competing jobs subject to limited resources. Second, the study of scheduling involves the theory of network, inventory, optimization and combinatorial problems which attract a great deal of active researchers. Third, the computer is available everywhere, which is an essential part to solve the scheduling problems.

REFERENCES

1. Ashour, S., "An Experimental Investigation and Comparative Evaluation of Flow Shop Scheduling Techniques", *Operations Research*, Vol. 18, No. 3, 1970.
2. Baker, K., "A Comparative Study of Flow-Shop Algorithms," *Operations Research*, Vol. 23, No. 1, 1975.
3. Baker, K., *Introduction to Sequencing and Scheduling*, John Wiley & Sons, Inc., New York, 1974.
4. Baker, K. and Martin, J., "An Experimental Comparison of Solution Algorithms for the Single-Machine Tardiness Problem", *Naval Research Logistics Quarterly*, Vol. 21, No. 1, 1974.
5. Balas, E., "Machine Sequencing: Disjunctive Graphs and Degree Contrained Subgraphs", *Naval Research Logistics Quarterly*, Vol. 17, No. 1, 1970.
6. Bennington, G. and McGinnis, L., "A Critique of Project Planning with Constrained Resources," in Symposium of the Theory of Scheduling and its Applications, S. Elmaghraby, Ed., Springer-Verlag, New York, 1973.
7. Burns, F. and Rooker, J., "A Special Case of the 3xn Flow Shop Problem", *Naval Research Logistics Quarterly*, Vol. 22, No. 4, 1975.
8. Campbell, H., Dudek, R. and Smith, M., "A Heuristic Algorithm for the n Job m Machine Sequencing Problem", *Management Science*, Vol. 16, No. 10, 1970.
9. Charlton, J. and Death, C., "A Generalized Machine Scheduling Algorithm", *Operational Research Quarterly*, Vol. 21, No. 1, 1970.
10. Conway, R., Maxwell, W. and Miller, L., *Theory of Scheduling*, Addison-Wesley Pub. lishing Company, Reading, Massachusetts, 1967.
11. Conway, R., "Priority Dispatching and Job Lateness in a Job Shop", *Journal of Industrial Engineering*, Vol. 16, No. 4, 1965.
12. Davis, E. "Project Network Summary Measures and Constrained Resource Scheduling," *AIIE Transactions*, Vol. 7, No. 2, 1975.
13. Davis, E. and Patterson, J. "A Compariosn of Heuristics and Optimal Solutions in Resource-constrained Project Scheduling," *Management Science*, Vol. 21, No. 8, 1975.
14. Dutta, S. and Cunningham, A. "Two-Machine Fow-Shops with Finite Intermediate Storage," *Management Science*, Vol. 21, No. 9, 1975.
15. Elmaghraby, S. "The Sequencing of Related Jobs," *Naval Research Logistics Quarterly*, Vol. 15, No. 1, 1968.
16. Florian, M., Trepant, P. and McMahon, G. "An Implicit Enumeration Algorithm for the Machine Sequencing Problem," *Management Science*, Vol. 17, No. 12, 1971.
17. Gelders, L. and Kleindorfer, P. "Coordinating Aggregate and Detailed Scheduling in the One-Machine Job Shop: II-Computation and Structure," *Operations Research*, Vol. 23, No. 2, 1975.
18. Gupta, J. "Optimal Schedules for Special Structure Flow Shops," *Naval Research Logistics Quarterly*, Vol. 22, No. 2, 1975.
19. Hershauer, J. and Ebert, R. "Search and Simulation of a Job-Shop Sequencing Rule," *Management Science*, Vol. 21, No. 7, 1975.
20. Holloway, C.and Nelson, R. "Job Shop Sc-

heduling with Due-date and Overtime Capacity," *Management Science*, Vol. 21, No. 1, 1974.

21. Holloway, C. and Nelson, R. "Job Shop Scheduling with Due-date and Variable Processing Times," *Management Science*, Vol. 20, No. 9, 1974.

22. Hottenstein, M. "Expediting in Job-Order-Control Systems: A Simulation Study," *AIIE Transactions*, Vol. 2, No. 1, 1970.

23. Karp, R. and Held, M. "Finite- State Processes and Dynamic Programming," *Journal of SIAM*, Vol. 10, No. 3, 1967.

24. Kolesar, P., Rider,K., Crabill, T. and Walker, W. "A Queuing-Linear Programming Approach to Scheduling Police Patrol Cars," *Operations Research*, Vol. 23, No. 6, 1975.

25. McMahon, G. and Florian, M. "On Scheduling with Ready Times and Due Dates to Minimize Maximum Lateness," *Operations Research*, Vol. 23, No. 3, 1975.

26. Moore, J. "An Algorithm fora Single Machine Scheduling Problem with Sequence Dependent Setup Times and scheduling Windows," *AIIE Transactions*, Vol. 7, No. 1, 1975.

27. Moore, J. and Wilson, R. "A Review of Simulation Research in Job Shop Scheduling," *Production and Inventory Management*, Vol. 8, No. 1, 1967.

28. Plamer, D. "Sequencing Jobs Through a Multi-Stage Process in the Minimum Total Time-A Quick Method of Obtaining a Near Optimum," *Operational Research Quarterly*, Vol. 16, No. 1, 1965.

29. Panwalkar, S. and Khan, A. "An Improved Branch and Bound Procedure for nxm Flow Shop Problems," *Naval Research Logistics Quarterly*, Vol. 22, No. 4. 1975.

30. Park, S. "A Branch and Bound Method for the Sequencing Problems of n Jobs on m Identical Machines in Parallel," *Unpublished M.Sc. thesis*, North Carolina State University, Raleigh, N.C., 1972.

31. Park, S. and Elmaghraby, S. "Scheduling Jobs on a Number of Identical Machines,: *AIIE Transactions*, Vol. 6, No. 1 1974.

32. Park, S. and Heiser, R. "On Single Machine Sequencing with Stochastic Processing Times and Stochastic Deferral Costs", *Presented at ORSA-TIMS Scheduling Conference*, Orlando, Florida, Feb. 4-6, 1976.

33. Rinnooy Kan, A., Lageweg, B. and Lenstra, J. "Minimizing Total Costs in One-Machine Scheduling," *Operations Research*, Vol. 23, No. 5, 1975.

34. Root, J. "Scheduling with Deadlines and Loss Functions on k Parallel Machines," *Management Science*, Vol. 11, No. 3, 1965.

35. Schrage, L. "Obtaining Optimal Solutions to Resource Constrained Network Scheduling Problems," *Proc. Systems Engineering Conference*, Phoenix, February 1971.

36. Shwimer, J. "On the n-Job, One-Machine, Sequence-Independent Scheduling Problem with Tardiness Penalties: A Branch and Bound Solution," *Management Science*, Vol. 18, No. 6, 1972.

37. Sidney, J. "Decomposition Algoriths for Single-Machine Sequencing with Precedence Relations and Deferral Costs," *Operations Research*, Vol. 23, No. 2, 1975.

38. Smith, M., Panwalkar, S. and Dudek, R. "Flowshop Sequencing Problem with Ordered Processing Time Matrices", *Management Science*, Vol. 21, No. 5, 1975.

39. Tilquin, C. "On Scheduling with Earliest Starts and Due Dates on a Group of Identical Machines," *Naval Research Logisitcs Quarterly*, Vol. 22, No. 4, 1975.

40. Uskup, E. and Smith S. "ABranch-and-Bound Algorithm for Two-Stage Production-Sequencing Problems," *Operations Research*, Vol. 23, No. 1, 1975.

41. Zaloom, V. and Vatz, D. "A Note on the Optimal Scheduling of Two Parallel Processors," *Naval Research Logistics Quarterly*, Vol. 22, No. 4, 1975.