DEPARTMENT OF OPERATIONS RESEARCH
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK

TECHNICAL REPORT NO. 116
CS 70-64

June 1970

SIMULATION OF THE JOB SHOP PROCESS

by

Richard W. Conway

Simulation of the Job Shop Process

Richard W. Conway

Cornell University

Discrete event simulation with a digital computer has been extremely important in the study of the scheduling procedures that might be used in a multi-machine job shop. It would probably not be an exaggeration to assert that all of the research results that have actually been applied in practice were obtained in this way, although the simulation studies were generally based upon analytical work. The simulation of job shops began in the mid-fifties at approximately the same time by groups at the University of California (Los Angeles), General Electric (Evendale) and IBM. It was one of the major areas of inquiry in the early period of digital simulation and it is apparent in any review of the early papers that the investigators were often learning as much about the technique of simulation as the intended subject area.

The first simulators were written in machine or assembly language and, although it was possible to produce a simulator in this way, it was slow and expensive and it soon became obvious that research by simulation requires frequent, rapid,and major changes in the model and these procedures could not accomodate such changes. The first attempts to improve the situation were the construction of generalized simulators to be specialized by the specification of parameters and the selection or provision of key subroutines. Such systems were produced by GE with IBM (Rowe, Rezucha), GE alone (Markowitz) and Cornell (Conway, Maxwell). None of these systems proved to have adequate flexibility and effort was soon directed to the development of special purpose higher-level computer languages. Three orginating in job shop work were SIMSCRIPT (RAND-Markowitz), CLP (Cornell-Conway, Maxwell) and GASP (US Steel-Kiviat). At about the same time SIMULA and GPSS appeared, The early versions of these languages exacted heavy penalties in computer time, but this was more than offset by their power and flexibility.

Another area in which existing tools proved inadequate for simulation was in the design of experiments and the statistical analysis of results. Standard statistical procedures were not directly applicable to this peculiar type of experimentation in which experiments could be precisely controlled, exactly duplicated, and produced time series rather than discrete observations. Some early investigators almost ignored questions of variability and significance. Today there is at least some recognition of these problems, but procedures are still very crude and both classical Fisherian statisticians and modern decision-theory statisticians would be appalled at what passes for statistical analysis of simulation results.

The most useful and general advice one can give with respect to measurement and estimation would be to only use simulation for direct comparison of alternative forms of the simulated system. Making comparisons between data produced by simulation and data generated by the corresponding real system is very difficult, and making absolute estimates is even harder. With some ingenuity the direct comparison approach can be adapted to most questions and is much to be preferred.

Two Types of Scheduling Simulation

There have been two quite different approaches to the scheduling problem, both using digital simulation. Some investigators have, in effect, been attempting to extend or generalize known theoretical results. A second group, faced with actual problems in real shops, have attempted to develop improved procedures under very specific and more or less realistic conditions. It is very difficult to assess the success that this second group has enjoyed since it is primarily industrial work and is rarely published. The conduct of the two types of studies is quite different and the practical, shop-modeling studies are comparatively difficult projects. The fraction of total effort required to collect and "discipline" data so that it is acceptable to the relatively delicate digestion of a simulation program is very high. The result is that such studies are long in preparation and often evaluate fewer scheduling alternatives than initially intended. The difficulty in preparing data is so great that one is impressed with the adaptive mechanisms that must exist in real shops--which digest such data without this pre-processing. and one can worry that perhaps the omission of these mechanisms in the model represents a serious departure from reality. The dimensions of real shops present another problem in the practical studies. Since the duration of a simulation run varies roughly with the square of the number of machines in the model, a one-to-one representation of a real shop often necessitates serious truncation of either the number or the length of experimental runs. The author has concluded that large dimensional modeling should be undertaken only after preliminary tests indicate that results are actually sensitive to dimension, and that real data collection should be undertaken only in desperation after very careful preliminary tests with data generated from extreme distribution forms has shown that real data are absolutely necessary.

In general, there seems to have evolved a feeling that simulation should err on the side of simplicity. It is far wiser to perform initial tests on a highly simplified form, and add detail, complexity and size only reluctantly and slowly. All of the early job shop models contained certain elegant complexities that were never used. These were constructed at considerable cost and effort, and although they might have contributed to the realism of the model had they been used, investigations did not reach the point where these features could be usefully employed during the lifetime of the original model. It is apparently axiomatic that simulation investigations proceed slowly and move in unexpected directions.

In terms of substantive results there do not appear to be serious discrepancies between the two approaches. The author and his colleagues, although primarily in the "theoretical" school, have participated in several practical studies, and discussed many others with their creators. There do not appear to be any great surprises--the practical studies are consistent with what might have been expected based on the more abstract and generalized work.

It is sometimes charged that simulation is a means for attaching quantitative estimates of dubious precision and validity to preconceived conclusions. Unfortunately this may well characterize some use, and simulation has undoubtedly ground many industrial axes. But however much abused it might be, it is nevertheless a powerful tool, and an experienced practitioner with a suitable language and adequate computing capacity can obtain significant insight into the operation of a complex system.

Multi-Machine Scheduling Results from Simulation Studies

It might be useful to illustrate some of the type of information about job shop scheduling procedures that simulation studies have produced. Much of the theoretical school of simulation has been devoted to trying to obtain an understanding of a simple multi-machine shop comparable to what analytical procedures have yielded for a single-machine shop (see Reference 4, Chapter I). Algebraic arguments for the static shop (in which all jobs are known and available at the beginning of the schedule period) and probabilistic arguments for the dynamic shop (in which the arrivals are a continuing stochastic process) have yielded considerable insight into the results that may be obtained from different sequencing procedures.. Alternatively, for many reasonable measures of performance optimal sequencing algorithms are known. However, analogous procedures and insight for multi-machine shops are not available and there is every indication that usefully computable procedures with optimal properties will not be found. Hence a number of simulation studies have attempted to obtain this insight experimentally, and have at least in part succeeded. The results have been interesting and useful, but clearly have yielded no optimal procedures.

A number of different investigations (References 1, 2, 3, 5, 6 and 7) have been performed under very similar conditions. These conditions describe a highly simplified job shop--actually a network of simple, single server queues. The principal restrictions are:

a) Each machine (server) is continuously available for assignment, without partition of the test period into shifts or days, and without intermittent unavailability (representing breakdown or rework).

b) Jobs are simple sequences of operations--no "assembly" operations are permitted.

c) Only one machine is capable of performing each particular operation.

d) No preemption is permitted--once the execution of a particular operation has begun it is carried to completion without interruption.

e) A job can be in process on at most one operation at any point in time--no overlap scheduling is permitted.

f) Each machine can handle at most one operation at a time.

g) Setup time is not separately identified and is independent of sequence.

h) Jobs are assumed to move instantaneously from one machine to the queue for the next.

These studies are essentially consistent in results for this simple shop, and the following extracts suggest the type of information that has been produced.

1) The average number of jobs in queue in a single-machine shop is known to be minimized by the shortest-processing-time discipline (SPT)--from jobs competing for assignment, select the one that has the minimum processing time. The attempt to generalize this result has been the central issue in much of the simulation study of the job shop. It is not intuitively obvious what the corresponding procedure would be in a multi-machine shop. One might select according to the minimum processing time for the imminent operation, or the minimum sum of all the processing times for all operations of the job, or the sum of the processing times for operations not yet performed on the job..Simulation has clearly established that tho first of these is the most advantageous form and that while this procedure does not possess the optimality in the multi-machine shop that it enjoys in the one-machine shop it nevertheless dominates the situation. For any measure of performance that involves average congestion in the shop this simple rule dominates all others--except for some composite rules that include major SPT influence. For example in one study (3) with nine machines

where each run represented 8700 jobs the following results were reported:

Scheduling procedure	Measure of Performance	
	Average number of jobs in shop	Average work remaining on all jobs in shop
Select minimum:		
Imminent operation processing time (SPT)	23.2	297
Total processing time on job	82.9	323
Remaining processing time on job	47.5	989
Select at random:	59.4	554
Select in order of arrival at machine:	58.9	559

Better performance than SPT was achieved only by complex rules that very often made the same selection as SPT, and differed only when the processing times were similar and the particular queue was not long.

2) The minimum variance of time-in-shop for individual jobs is achieved by ordering the jobs at a particular machine according to the time at which they arrived at the shop (not time arrived at that particular machine). Rather surprisingly, SPT is second only to this arrival ordering in producing small shop time variance.

3) In a single-machine shop it is known that sequencing jobs by due-date minimizes the maximum lateness (lateness being the difference between completion-date and due-date); sequencing by SPT minimizes the average lateness; and sequencing by minimum "slack-time" (due date minus processing time) maximizes the minimum lateness. It is not known how to minimize the number of jobs late, or the average tardiness (tardiness is MAX(0,lateness)). Multi-machine results are strikingly different. From the same study cited above (3) the following results were reported:

Scheduling procedure	Measure of Performance		
	Average time in shop	Average lateness	Percentage with lateness > 0
Select in order of arrival at machine	74.4	-4.5	44.8%
Select earliest due-date	63.7	-15.5	17.7
Select minimum "slack-time"	65.8	-13.1	22.0
Select minimum: $\frac{\text{slack-time}}{\text{no. opns remaining}}$	66.1	-12.8	3.7
Select minimum operation proc time (SPT)	34.0	-44.9	5.0

These results led to procedures that combined an SPT influence with consideratio of the slack-time per operation. The relative importance of these two influence was dynamic, depending upon the backlog of work at the particular machine for which the selection was to be made: with large backlog the SPT influence was dominant; with small backlog the SPT influence was minimal. The following results for such a procedure were reported by a later study (5) using a set of 6000 jobs with much more difficult due-dates:

Scheduling procedure	Measure or Performance			
	Average lateness	Percent tardy	Average tardiness All jobs	Tardy jobs
Select earliest due-date	18.2	49.8	35.2	70.7
Select minimum operation processing time (SPT)	-49.2	10.8	12.1	112.1
Select minimum slack per operation remaining	-1.2	42.2	21.9	51.9
Select by dynamic combination	-39.4	12.0	5.4	45.3

This provides a good example of the use of simulation to develop progressively more powerful procedures. It also illustrates the not-too-common occurence of completely unexpected results. The combination procedure somehow extracts the best characteristics of each of its components, and does not behave as a compromise between the two.

4) The relative performance of sequencing procedures appears to be relatively insensitive to the size of the shop and the distribution of processing times. The relative advantage of SPT sequencing is persistent and substantial. Even when decisions must be based on estimates of processing time (since the actual time is not known until processing is completed) SPT dominates and is surprisingly insensitive to errors of estimate.

5) The one condition that is crucial is the prohibition of assembly operations. Preliminary work on this question indicates that it is basically a different model and that entirely different procedures become important.

The Future Role of Simulation

To date simulation has essentially been a research and development tool. It should eventually assume a quite different role with respect to job shop scheduling, and perhaps in some few pioneering installations is already doing so. This role involves becoming an integral part of an actual production control system. Computing systems are already used to maintain current status of shops in real-time, and to generate job assignment decisions by means of pre-established decision rules (which may be the result of an investigation involving simulation.) What remains is to construct the program for the production control system to have two modes of operation. The normal mode would be similar to what is now done with such systems. The alternate, or simulation, mode would permit the system to "run ahead" for a short period in order to predict at least the short-run effect of various choices that might be made in an imminent decision. Presumably this simulated run-ahead could be repeated or extended until the controller was satisfied as to the best choice. The system would then revert to a more prosaic status-keeping role until the next decision point was reached. It would not seem that such a system would be enormously more difficult to implement than the existing computer production control systems--certainly not more than a factor of two. If it yielded only minor improvement over the performance of a conventional status-keeping, decision-rule system it would quickly recover this construction cost.

References:

1. W. Bakker, "De Levertijd in Machinefabrieken," thesis, Technische Hogeschool at Delft, Holland, June 1965
2. D. C. Carroll, "Heuristic Sequencing of Single and Multiple Component Jobs," Ph.D. thesis, MIT, June 1965
3. R. W. Conway, "An Experimental Investigation of Priority Assignment in a Job Shop," RAND Corporation Memorandum RM-3789-PR, February 1964
4. R. W. Conway, W. L. Maxwell, L. W. Miller, "Theory of Scheduling," Addison-Wesley Publishing Co., Reading, Mass. 1967.
5. J. W. Oldziey, "Dispatching Rules and Job Tardiness in a Simulated Job Shop," thesis, Cornell University, February 1966
6. R. T. Nelson, "Labor and Machine Limited Production Systems," Management Sciences Research Project, UCLA, Paper 90, January 1966
7. Y. R. Nanot, "An Experimental Investigation and Comparative Evaluation of Priority Disciplines in Job-Shop-Like Queuing Networks," Management Sciences Research Project, UCLA, Research Report 87, December 1963.