Engineering Costs and Production Economics, 18 (1990) 293-299
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

MAXIMIZING THROUGHPUT IN SOME SIMPLE TIME-CONSTRAINED SCHEDULING SITUATIONS*

H.H. Huisman, G.L. Polderman

Philips International B.V., Eindhoven (The Netherlands)

and P.J. Weeda

University of Twente, Enschede (The Netherlands)

ABSTRACT

This contribution focuses on some interesting phenomena occurring in scheduling capacity constrained resources in time-constrained situations. The scheduling situations considered form part of a simulation game developed to assist in teaching the scheduling philosophy of Optimized Production Technology (OPT) to production managers. It is shown that under certain conditions on set-up and processing times, advanced examples may be constructed combining three of the four complicating conditions mentioned in the OPT-literature. In addition, some interesting properties of process batches yielding a maximum throughput in such cases is considered.

1. INTRODUCTION

In spite of its mysteries, an increasing number of corporations have purchased and are using the production control software package Optimized Production Control (OPT) originally developed by Creative Output Inc. Obviously, the OPT philosophy has sufficient appeal to convince topmanagers in the U.S. as well as in Europe about its potential for improving throughput. The secret nature of one of its main operational modules has not been considered as a serious barrier. This fact justifies the conclusion that the simulations executed with the OPT-system in these corporations could stand the comparison with reality. For the history of the development of OPT-software and its philosophy the reader is referred to [1,2].

In The Netherlands, Philips International B.V. is presently doing pioneering work on the implementation of the OPT-software in several of its factories. This development created the need to familiarize production employees with the scheduling philosophy of OPT. For this purpose a simulation game has been developed. Such a game is also useful if the scheduling philosophy of OPT is used without actually buying the OPT-software. A critical evaluation of the OPT-scheduling system is presented in [3].

Experiments with the game can be performed on (within certain limits on size) ar-

*Presented at the 5th International Working Seminar on Production Economics, Igls, Austria, February 22-26, 1988.

0167-188X/90/$03.50

bitrary process configurations. A process configuration is defined by a number of processes, which are run through by products in accordance with a specific routing. To each process a unique machine is assigned. Machines may be assigned to several processes. Each pair of process and machine implies a set-up time and a unit processing time. The size and the timing of transfer batches are at choice. A given time allowance T is specified during which throughput should be maximized. The task is to determine how batch sizes should be chosen in order to meet this criterion. An alternative approach to the study of interactions between batch sizing and system performance is developed in [4].

In this paper a particular process configuration with a serial structure consisting of three processes and two machines is analyzed mathematically. The results were used to provide hints for the user of the game. Moreover it is shown that this configuration provides for advanced examples on certain aspects of OPT. In OPT-terminology a capacity constrained resource is defined as any resource that will disrupt the flow if not taken into consideration when scheduling. Four conditions complicating the scheduling are mentioned in the OPT-literature (cf. [5]). These four conditions are:

(1) Different lead times from capacity constrained resource to due dates
(2) Set-up on a capacity constrained resource
(3) One capacity constrained resource feeding more than one part to the same product
(4) One capacity constrained resource feeding another one.

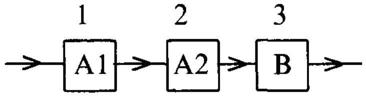
Condition (1) means that remaining slack times of the various jobs are different. Remaining slack time is calculated as the time interval between the current time and the due data minus the remaining processing time. This condition cannot be created by the configuration discussed here. However, combinations of the remaining three conditions are illustrated below. Occasionally some of the nine socalled OPT–rules (cf. [6,7]) are brought into the discussion. These nine rules are:

1. Balance flow, not capacity
2. Constraints determine non-bottleneck utilization
3. Activation is not always equal to utilization
4. An hour lost on a bottleneck is an hour lost for the whole system
5. An hour saved at a non-bottleneck is a mirage
6. Bottlenecks govern throughput and inventory
7. A transfer batch should not always equal a process batch
8. A process batch should be variable, not fixed
9. Set the schedule by examining all constraints simultaneously.

In this paper only machines can be capacity constraints and the term critical capacity will be used. A machine is a critical capacity if it determines the maximum throughput of the process configuration. A critical capacity is fully occupied except for an inevitable idle time due to the finiteness of the allowed time interval.

2. ANALYSIS OF THE PROCESS CONFIGURATION

The configuration under consideration is depicted below.



A raw material is successively subject to three manufacturing processes 1, 2 and 3. Process 1 and process 2 are executed in this order on machine A and process 3 is executed on machine B. The set-up times are respectively denoted by $S(\mathrm{A1})$, $S(\mathrm{A2})$ and $S(\mathrm{B})$. The simplicity $S(\mathrm{B})$ is taken equal to zero. The processing times per unit product are respectively denoted by $P(\mathrm{A1})$, $P(\mathrm{A2})$ and $P(\mathrm{B})$. Frequently the notations $S(\mathrm{A})=S(\mathrm{A1})+S(\mathrm{A2})$ and $P(\mathrm{A})$

$=P(\mathrm{A1})+P(\mathrm{A2})$ are used. Each unit product is immediately transferred to the next process, i.e. transfer batches are equal to one.

A detailed analysis of this process configuration requires a subdivision with respect to different relations between the unit processing times. Primarily the set-up times on machine A are assumed to vanish, i.e. $S(\mathrm{A1})=S(\mathrm{A2})=0$. Two cases need to be distinguished. Afterwards both set-up times on machine A are assumed to be positive and three cases need to be distinguished.

Case 1. $S(\mathrm{A1})=S(\mathrm{A2})=0$ and $P(\mathrm{A})>P(\mathrm{B})$

In this case process 1 and process 2 can be considered as one process with unit processing time $P(\mathrm{A})$. Each unit product processed on machine B is ready before the next product on machine A is finished. Machine A produces uninterruptedly and is only idle at the end of time interval T. During that idle time the last operation on B is processed. The maximum throughput $M\mathrm{A}(T)$ is given by

$$M\mathrm{A}(T)=\frac{T-P(\mathrm{B})}{P(\mathrm{A})} \tag{1}$$

while machine A is the critical capacity. Note that this situation satisfies complication condition (3): One capacity constrained resource (machine A) feeding more than one part to the same product (process 3).

Case 2. $S(\mathrm{A1})=S(\mathrm{A2})=0$ with $P(\mathrm{A})<P(\mathrm{B})$

In this case a product unit processed on machine A is ready before the preceding product unit on machine B is finished. Machine B is only idle during the processing time of the first product on machine A. The maximum throughput $M\mathrm{B}(T)$ is given by

$$M\mathrm{B}(\mathrm{T})=\frac{T-P(\mathrm{A})}{P(\mathrm{B})} \tag{2}$$

and machine B is the critical capacity.

Case 3. $S(\mathrm{A1}), S(\mathrm{A2})>0$ and $P(\mathrm{A2})>P(\mathrm{B})$

In this case at least two complicating conditions are satisfied: "Set-up on a capacity constrained resource" and "One capacity constrained resource feeding more than one part to the same product". Since set-up times on machine A are positive, process 2 on machine A cannot be started before the batch of process 1 on A is finished. Process 3 on machine B can get started as soon as one product unit has gone through both processes on A. Since $P(\mathrm{A2})>P(\mathrm{B})$, machine A should produce uninterruptedly. It will only be idle at the end of the allowed time interval T when the last product unit on machine B is processed. To save set-up time a single batch should be processed. Its size equals the maximum throughput given by

$$M\mathrm{A}(T)=\frac{T-S(\mathrm{A})-P(\mathrm{B})}{P(\mathrm{A})} \tag{3}$$

while machine A is the critical capacity.

Case 4. $S(\mathrm{A1}), S(\mathrm{A2})>0$ with $P(\mathrm{A})<P(\mathrm{B})$

In this case to the number of complicating conditions a third can be added: "One capacity constrained resource feeding another one". Since $P(\mathrm{B})>P(\mathrm{A})$ one would assume that machine B is the critical capacity. Adopting this point of view machine B should process uninterruptedly except for an idle time at the start of the interval T in which machine B is waiting for the first product unit finished with process 2 on machine A. The maximum throughput under this assumption, $M\mathrm{B}(T)$, is given by

$$\mathrm{MB}(T)=\frac{T-S(\mathrm{A})-QP(\mathrm{A1})-P(\mathrm{A2})}{P(\mathrm{B})} \tag{4}$$

where Q denotes the size of the first batch.

So far, this procedure does not specify the number of batches on A nor their sizes. To determine these, observe that during the processing of a batch on machine B process 2 still has to be finished on machine A. Only thereafter a second batch can be started on machine A. Machine B can start again with the second batch as soon as or after the first unit product of process 2 of the second batch is finished on machine A. Hence in order to guarantee uninterrupted processing on machine B two successive batches with sizes $Q1$ and $Q2$ should satisfy

$$Q1P(\mathrm{B}) \geqslant (Q1-1)P(\mathrm{A2})+S(\mathrm{A})+Q2P(\mathrm{A1})+P(\mathrm{A2}) \tag{5}$$

or equivalently

$$Q1P(\mathrm{B}) \geqslant Q1P(\mathrm{A2})+S(\mathrm{A})+Q2\dot{P}(\mathrm{A1}) \tag{6}$$

If batch sizes are repetitive, i.e. $Q1=Q2=Q$, Q should satisfy

$$QP(\mathrm{B}) \geqslant QP(\mathrm{A2})+S(\mathrm{A})+QP(\mathrm{A1}) \tag{7}$$

or equivalently

$$Q \geqslant \frac{S(\mathrm{A})}{P(\mathrm{B})-P(\mathrm{A})}=Q' \tag{8}$$

The right hand side of (8) specifies a batch size Q', for which the equality sign holds in (7). Q' can be interpreted as the batch size for which the processing times including set-up time are equal for both machines. This is surprising since the consequence of a repetitive batch size Q' is that both machines are uninterruptedly occupied. Hence both machines are in fact critical capacities.

It can be argued that in case B is the critical machine, the best choice for Q is indeed $Q=Q'$. If $Q<Q'$ is chosen then (7) is violated and machine B has inter-batch waiting times. If $Q>Q'$ then the idle time of machine B at the start of the interval T is extended. Formula (4) shows that in this case the throughput decreases. If $Q<Q'$ machine A is uninterruptedly occupied an hence determines the throughput, given by

$$MA(T)=\frac{T-QP(\mathrm{B})+(Q-1)P(\mathrm{A2})}{[S(\mathrm{A})+QP(\mathrm{A})]/\mathrm{Q}} \tag{9}$$

where $QP(\mathrm{B})-(Q-1)P(\mathrm{A2})$ represents the idle time on machine A at the end of time interval T.

By comparing (4) and (9) it is easily verified that for $Q=Q'$ or equivalently $S(\mathrm{A})+QP(\mathrm{A})=QP(\mathrm{B})$ the maximum throughput is given by

$$MA(T)=MB(T) \tag{10}$$

and is obtained for a repetitive batch size equal to Q'. The case that varying batch sizes are permitted will not be considered here. For $Q<Q'$ machine A governs the throughput which diminishes if Q decreases, since the total set-up time increases. For $Q>Q'$ machine B governs the throughput, which diminishes if Q increases as already remarked above.

Summarizing the case that the unit processing time satisfy the relation $P(\mathrm{B})>P(\mathrm{A1})+P(\mathrm{A2})$ there exists a critical batch size Q' for which both machines are uninterruptedly and equally utilized and the maximum throughput is obtained. In relation to OPT-rule 1 this means that flow as well as capacity is balanced for this critical batch size: balance in optima forma! Furthermore this example illustrates the OPT-rule "Set the schedule by examining all constraints simultaneously".

Case 5. $S(\mathrm{A1})$, $S(\mathrm{A2})>0$ with $P(\mathrm{A2})<P(\mathrm{B})<P(\mathrm{A})$

Because $P(\mathrm{A})>P(\mathrm{B})$, one would regard machine A as the critical capacity. To save set-up time a single batch should be considered with size Q satisfying

$$T=S(\mathrm{A})+QP(\mathrm{A1})+P(\mathrm{A2})+QP(\mathrm{B}) \tag{11}$$

or solving for Q

$$Q=\frac{T-S(\mathrm{A})-P(\mathrm{A2})}{P(\mathrm{A1})+P(\mathrm{B})} \tag{12}$$

However, since $P(\mathrm{B})>P(\mathrm{A2})$ a second set-up on machine A becomes possible. This would reduce the inter-batch waiting time of machine B until the first product of process 2 of the second batch on machine A is ready. In fact machine B is then considered as critical capacity. Suppose that a second set-up of machine A is considered with batch size $Q2$ then machine B would be able to process uninterruptedly if relation (6) holds. Note that in this case relation

(8) would result in a critical batch size $Q' < 0$, since $P(\mathrm{B}) < P(\mathrm{A})$. On the other hand $P(\mathrm{B}) \leqslant P(\mathrm{A})$ and $S(\mathrm{A}) > 0$ imply for batch size $Q1$

$$Q1P(\mathrm{B}) < S(\mathrm{A}) + Q1P(\mathrm{A}) \tag{13}$$

or equivalently

$$Q1[P(\mathrm{B}) - P(\mathrm{A2})] < S(\mathrm{A}) + Q1P((\mathrm{A1}) \tag{14}$$

Relation (6) can be written in the equivalent form

$$Q1[P(\mathrm{B}) - P(\mathrm{A2})] \geqslant S(\mathrm{A}) + Q2P(\mathrm{A1}) \tag{15}$$

Combining relations (14) and (15) yields

$$S(\mathrm{A}) + Q1P(\mathrm{A1}) > Q1[P(\mathrm{B}) - P(\mathrm{A2})] \geqslant S(\mathrm{A}) + Q2P(\mathrm{A1}) \tag{16}$$

which implies $Q1 > Q2$. Hence it is impossible to keep machine B processing unless the second batch size $Q2$ satisfies: $Q2 < Q1$. By introduction it follows that machine B can only be considered as the critical capacity if successive batch sizes form a decreasing sequence. This suggests the following procedure. Assume a starting batch size equal to $Q(1)$ and determine the sequence $Q(2), Q(3), \ldots$ such that

$$Q(n)P(\mathrm{B}) \geqslant Q(n)P(\mathrm{A2}) + S(\mathrm{A}) + Q(n+1)P(\mathrm{A1}) \quad n = 1,2\ldots \tag{17}$$

or equivalently

$$Q(n+1) \leqslant \frac{Q(n)[P(\mathrm{B}) - P(\mathrm{A2})] - S(\mathrm{A})}{P(\mathrm{A1})} \quad n = 1,2\ldots \tag{18}$$

The sequence is finite since ultimately the minimum batch size of one is understepped. Because $Q(1)$ is unknown, a feasible computational procedure would be to start with batch size one and work backwards by computing $Q(n)$ from $Q(n+1)$ in relation (18) with the equality sign until the allowed time T is overstepped. Next one starts with an initial batch size of two and repeats the computation and so on. Ultimately the sequence with the maximum throughput is obtained on the premise that B is the critical machine and therefore fully occupied. The backward recursive relation is given by

$$Q(n-1) = \frac{Q(n)P(\mathrm{A1}) + S(\mathrm{A})}{P(\mathrm{B}) - P(\mathrm{A2})} \tag{19}$$

If $Q(n-1)$ is non-integer then it should rounded off upwards, since (19) is in fact a larger than or equal condition.

The procedure will be illustrated on the following example:

$S(\mathrm{A1}) = S(\mathrm{A2}) = 70 \quad P(\mathrm{A1}) = P(\mathrm{A2}) = 30$

$P(\mathrm{B}) = 50$

$T = 2400$ (= 40 h)

$Q(n) = 1 \quad Q(n-1) = 8.5 \rightarrow 9$

$\mathrm{time} = S(\mathrm{A}) + Q(n-1)P(\mathrm{A1}) + P(\mathrm{A2}) + [Q(n-1) + Q(n)]P(\mathrm{B}) = 940$

$Q(n-1) = 9 \quad Q(n-2) = 20.5 \rightarrow 21$

$\mathrm{time} = S(\mathrm{A}) + Q(n-2)P(\mathrm{A1}) + P(\mathrm{A2}) + [Q(n-2) + Q(n-1) + Q(n)]P(\mathrm{B}) = 2350$

$Q(n) = 2 \quad Q(n-1) = 10 \quad Q(n-2) = 22$

time = 2530 > 2400 → solution 1 is feasible

The batch times and sizes of the machines are:

machine A	machine B	batch size
(0,1400)	(800,1850)	21
(1400,2080)	(1850,2300)	9
(2080,2280)	(2300,2350)	1

Note that both machines are processing uninterruptedly.

This example demonstrates also the three complicating conditions already mentioned. Furthermore it underlines the OPT-rule "Process batches should be variable, not fixed", since uninterrupted utilization of machine B implies strictly decreasing batch sizes. If machine A is assumed to be the critical capacity and a single batch of size Q is processed, then according to (12), Q is given by

$$Q=\frac{2400-140-30}{30+50}=27.9$$

Obviously this assumption does not maximize throughput.

If in this example the allowed time is increased by 10, then this procedure would yield the same results. However, the following solution shows that allowing an inter-batch waiting time of 10 between the processing of the second and the third batch on machine B, the throughput becomes 32 instead of 31. In this case only machine A is processing without interruptions, as the results below show.

machine A	machine B	batch size
(0,1400)	(800,1850)	21
(1400,2080)	(1850,2300)	9
(2080,2340)	(2310,2410)	2

This counterexample shows that the maximum throughput is not guaranteed by the procedure outlined above. It has been based on the assumption of machine B being the critical machine. Hence this assumption is neither a sufficient nor a necessary condition for maximum throughput, since in the counter example machine A takes over the role of critical machine B. Furthermore it shows that uninterrupted utilization of both machines is, although sufficient, not a necessary condition for maximum throughput. In any case it illustrates again the OPT-rule "Set the schedule by examining all constraints simultaneously".

3. CONCLUSIONS

In Section 2 a process configuration has been analysed in order to demonstrate on advanced examples combinations of three conditions, which complicate the scheduling of capacity constrained resources according to the philosophy of Optimized Production Technology (OPT). In addition properties of schedules maximizing throughput have been discussed. The process configuration analyzed consists of two processes (on machine A) feeding into a third process (on machine B). Its structure incorporates the complicating condition "One capacity constraint feeding more than one part to the same product", only if machine A is the capacity constraint.

According to zero or positive set-up times and different relations between unit processing times, five cases have been considered. They will be summarized successively.

Case 1. $S(\mathrm{A1})=S(\mathrm{A2})=0$ and $P(\mathrm{A1})+P(\mathrm{A2})>P(\mathrm{B})$

Machine A is the critical capacity. Machine A is fully utilized during the allowed time interval T except for an idle time at the end of T, which is equal to the processing time of the last unit on machine B.

Case 2. $S(\mathrm{A1})=S(\mathrm{A2})=0$ and $P(\mathrm{A})<P(\mathrm{B})$

Compared with case 1 the roles of the machines are reversed. Machine B is the critical capacity. It is fully utilized during the allowed time interval T, except for an idle time at the start, which is equal to the completion time of the first unit on machine A.

Case 3. $S(\mathrm{A1})$, $S(\mathrm{A2})>0$ and $P(\mathrm{A2})>P(\mathrm{B})$

In this case complicating condition: "Set-up on a capacity constraint" is added to the one already valid. Machine A is the critical capacity. Similar to case 1 a single batch is produced, which fully utilizes machine A. The only difference is that set-up time is included in the batch processing time.

Case 4. $S(\mathrm{A1}), S(\mathrm{A2})>0$ and $P(\mathrm{A})<P(\mathrm{B})$

In this case complicating condition: "One capacity constrained resource feeding another one" is added to the two already valid. There exists an optimal repetitive batch size. For this batch size both machines are uninterruptedly utilized. They are also equally utilized since batch processing times are equal on both machines. This critical batch size appears to be independent of the time constraint. This case also illustrates the OPT-rule: "Set the schedule by examining all constraints simultaneously".

Moreover flow as well as capacity is balanced in relation to OPT-rule 1.

Case 5. $S(\mathrm{A})>0$ and $P(\mathrm{A}2)<P(\mathrm{B})<P(\mathrm{A})$

In this case all three complicating conditions are valid. However, the critical batch size is negative. If machine A is assumed to be the critical capacity and a single batch is processed, then throughput is not maximized. If machine B is assumed to be the critical capacity then it is uninterruptedly utilized for a strictly decreasing sequence of batch sizes. This property is independent of the time constraint. It confirms the OPT-rule "Process batches should be variable, not fixed". However, uninterrupted utilization of machine B is neither a sufficient nor a necessary condition for maximum throughput, as a counter example reveals. Uninterrupted utilization of both machines will be a sufficient, but certainly not a necessary condition for maximum throughput.

In any case the OPT-rule: "Set the schedule by examining all constraints simultaneously" is valid.

In conclusion it is remarked that all conclusions of this paper are immediately extendable to an arbitrary number of processes on machine A. Suppose there are n processes A1, A2, ..., An on machine A. All formulas remain valid if $P(\mathrm{A}2)$ is replaced by $P(\mathrm{A}n)$ and $P(\mathrm{A}1)$ by the sum of the first $(n-1)$ unit processing times on A. $P(\mathrm{A})$ and $S(\mathrm{A})$ have to be interpreted respectively as the sum of the n unit processing times and $S(\mathrm{A})$ as the sum of the n set-up times.

REFERENCES

1 Goldratt, E.M., 1988. Computerized shop floor scheduling. Int. J. Prod. Res., 26: 443–455.

2 Goldratt, E.M. and Cox, J., 1984. The Goal, Excellence in Manufacturing. North River Press, New York.

3 Jacobs, F.R., 1983. The OPT scheduling system: A review of a new production scheduling system. Prod. Inv. Manage., 24(3): 47–51.

4 Jacobs, F.R. and Bragg, D.J., 1988. Repetitive lots: Flowtime reductions through sequencing and dynamic batch sizing. Dec. Sci., 19(2): 281–294.

5 Goldratt, E.M. and Fox, R.E., 1986. The Race. North River Press, New York.

6 Jacobs, F.R., 1984. OPT uncovered: Many production planning and scheduling concepts can be applied with or without software. Ind. Eng., 16(10): 32–41.

7 Lundrigan, R., 1986. What is this thing called OPT? Prod. Inv. Manage., 27(2): 2–12.

(Accepted July 8, 1989)