

An artificial, intelligence-based production scheduler

Alberto De Toni, Guido Nassimbeni and Stefano Tonchia Department of Electrical, Managerial and Mechanized Engineering (DIEGM), University of Udine, Italy

Describes a production scheduler, which utilizes a hybrid push/pull approach to schedule and exploits the expert system technology in order to obtain satisfactory solutions. The scheduler is applied to a multi-stage production and inventory system, managed by make-to-order, with a large variety of incoming orders. The search for solution is made in respect of the due-dates and under efficiency constraints (minimum lot, maximum storehouse levels, etc.). Considers order aggregation, both at portfolio and production level. Provides a dynamic rescheduling mechanism. Outlines theoretical arguments in favour of the scheduler and notes practical advantages as a consequence of the application of the scheduler in a firm which utilized a traditional despatching system.

Introduction

Intelligent solutions, based on expert systems, to solve problems of equipment diagnosis, process control and system design are becoming more and more widespread. The field of production scheduling has also been explored, but few real applications exist, perhaps because today's expert system technology needs a precise focus in developing an application, and production scheduling problems and approaches are often too general. The authors have developed an expert system for scheduling in make-to-order firms: in this paper they present the results obtained when applying this system. This expert system is a production scheduler with an original model-based reasoning. It can be considered a possible solution to the problem of scheduling orders along a production line with a due-date for the assembly of the final product and a finite capacity in the centres which constitute the production line.

The main field of applicability is that of make-to-order firms, where a trade-off exists between customer service and production efficiency:

- customer service, in terms of respecting due-dates and permitting a small order quantity (i.e. offering variety);
- production efficiency, in terms of working lots of a large size (i.e. with rare machine set-ups).

The queues before each work centre are the critical elements which must be controlled. Experience has demonstrated that an acceptable solution can be arrived at from a correct level in the queues and their stability. Queues must be kept stable in order to guarantee an adequate backlog (i.e. work load) for each work centre (no centre should ever find itself starved) and, at the same time, avoid an excessive delay to an item at each work centre.

According to the proposed reasoning model, queues are a dynamic fact and are managed by production-order release criteria and customer-order aggregation criteria.

The aggregation criteria are formulated in order to compact the orders of several customers to obtain a minimum lot size for production. Customer-order aggregation

criteria, usually only at portfolio level, permit the choice of the releases on the basis of the due-date and of the type of work requested by each single order. Production-order release criteria are based on the minimum slack time priority rule, and consider the inventory levels.

The problem

One of the typical situations of lot production is represented by several work centres arranged along production lines, with final assembly centres: stores are placed between the work centres, to cope with the asynchronous flow arising from different working parameters in each centre. This configuration is often called multi-stage production.

The case examined as an example consists of three lines and two final assembly centres (Figure 1). Each centre is characterized by three different working parameters:

- 1 minimum lot size;
- 2 set-up time;
- 3 run time per unit.

Each customer order, consisting of two specifications (type and quantity), determines a production order for each work centre.

Working with a minimum lot size means that all the production order releases must be greater than or equal to that size, even if the requirements were less. For example there may be a need for only 80 pieces, but with a minimum lot size of 100, 100 pieces are released, while with a request for 120 pieces, 120 pieces are released. The minimum lot size is a constraint related to production efficiency, in order to reduce the total time spent in set-up activities.

The complexity in managing a multi-stage production system is due to the different optimal lot sizes for each centre, further complicated by different set-up and run times between the centres.

The artificial intelligence-based approach can offer advantages, in respect of a traditional algorithmic (or deterministic) solution. This can be summarized as follows:

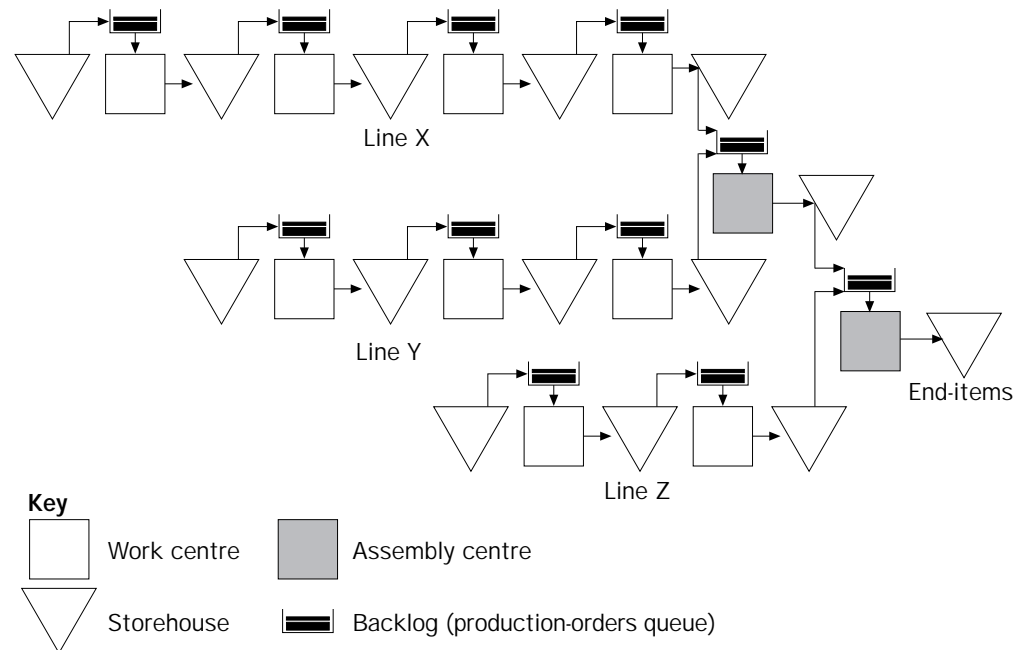
- There are rules for reaching the objectives (such as respecting due-dates and working

A first limited version of this work was presented at the 30th MATADOR Conference, UMIST, Manchester, 31 March-1 April 1993, and included in the *Proceedings* (A.K. Kochlar, Ed.).

Integrated Manufacturing Systems
7/3 [1996] 17-25

© MCB University Press
[ISSN 0957-6061]

Figure 1
Multi-stage production and inventory



with lots of large size) and rules to manage the priorities of the objectives.

- Restrictions can be made in the space of solution by means of heuristic rules, or constraints can be opportunely relaxed if no solution is found (in production scheduling problems, the space of solution is typically very wide, connected to several variables which can be dealt with – levels of semi-finished goods in storehouses, levels of backlogs, jobs actually being processed, etc.).
- Situations can be evaluated because of the separation between the descriptive part and the procedural part in the knowledge base.

The descriptive part (or base of facts) describes the status of the order portfolio, the actual status of each work centre and its backlog, the status of intermediary and final storehouses, and the working parameters of each work centre.

The procedural part (or base of rules) contains the rules which must translate the decisional process performed by a human expert in solving the problem of the production order release into a reasoning model respecting the constraints:

- completion of all the orders within their due-date;
- processing only one production order at a time on each machine of each work centre;
- finite capacity of each work centre;
- maximum and minimum intermediary storehouse levels;

- maximum number of machine set-ups per period;
- maximum and minimum queue length (in time unit) before each work centre.

Description of the artificial intelligence-based production scheduler

The reasoning model, translated into rules of the IF/THEN type, permits the construction of a lot production scheduler able to release, at the appropriate moment, lots of satisfactory dimension in respect to the predefined constraints (with levels derived from experience).

The developed intelligent production scheduler works with an inferential engine in forward chaining: the engine matches all the assertions of the IF part of a rule and, if they are all present in the base of facts, enriches the base of facts itself with the assertions contained in the THEN part of the rule. For example:

```
(DEFINE-RULE centre_setup (:PRIORITY 200))
(present_centre ?upper_centre)
(INSTANCE ?upper_centre IS centre WITH
lower_centre ?lower_centre)
(centre_status ?lower_centre waiting)
(centre_status ?upper_centre free)
(material_for ?upper_centre available)
THEN
```

(setup ?upper_centre
for_the_process_required_by ?lower_cen-
tre))

In the example the terms preceded by “?” are the variables: the variable “?upper_centre” acquires the name of the second term in the list whose first term is “present_centre”.

Then the name of the lower centre (it is the variable “?lower_centre”) is singled out. If this is waiting, if its upper centre is free and if there is material to process, then the list is determined which schedules the set-up of the upper centre as required by the lower centre.

Now we will describe the reasoning model.

A discrete-event system

First of all, it must be said that the productive context is considered by the production scheduler as a discrete-event system: the status of orders, work centres, queues and storehouses is periodically modified owing to events such as the release of an order, a change in machine set-up, etc. This temporization in single events, in addition to reflecting the functioning of the productive context, allows simulations to be realized: in fact, during the interval between one event and the next, a reconsideration and re-evaluation of already released but still not executed orders can be made.

Scheduling determines the starting and finishing date of each production order at each work centre in relation to each released customer order. So the problem is how aggregations of customer orders can be made and which customer orders aggregation must be released to satisfy the different optimum lot sizes requested by the different work centres along the production lines and at the final assembly stages.

The reasoning model is based on two fundamental steps:

- 1 defining the customer orders aggregations;
- 2 scheduling the production orders at each work centre, in relation to the decision regarding the customer orders aggregation release.

The customer orders aggregation

The rules which start off the aggregation mechanism involve a search for customer orders with the same product-code and more or less similar due-dates (a predetermined maximum time interval exists). The aggregation of the customer orders must form production orders in each single work centre which are greater than the minimum lot size characteristic of each work centre.

If the possible aggregations do not reach the minimum lot size, these aggregations can

be released into production in any case, by adding units so as to reach the minimum lot size. We call these free units, they are not produced to meet a specific order and they constitute a supplementary load.

The production orders scheduling

The reasoning process which, in the simulation mode, leads to the real lot production scheduling can be summarized as follows:

- loading of the work centres in simulation mode;
- calculation of the slack time for all the customer orders aggregations;
- customer orders aggregation selection for the real release;
- resources allocation and production order releases (one for each work centre).

For each customer orders aggregation, a work centre loading is simulated as a consequence of the release of the customer orders of that aggregation, in the form of production orders at each work centre:

- Each customer orders aggregation in the portfolio induces a search in the storehouse for free units of finished products and successively among the free units in the backlog of the immediately preceding centre, with the aim of filling these orders.
- If after this search there is still a definite requirement, a release in that centre of the production line is simulated, after the production orders already in the queue.
- This release must take place according to the minimum lot size practice and the work time of the centre is calculated as unit time by lot size (greater than or equal to the minimum lot size) plus the time for machine set-up, if needed.
- The simulation of the release in this work centre leads to a search for free units of material to process in the upstream storehouse and eventually among the free units in the backlog of the centre immediately preceding it.
- This logic moves backwards along the production system, as far as the raw material storehouses, obtaining a simulated allocation of the materials and a simulated loading of the work centres for the release of these orders.

For each customer orders aggregation, the slack time is calculated as the difference between the due-date and the throughput time needed to complete those orders. If this value is negative it means that it is impossible to complete the orders in time and it constitutes a measure of delay. The customer orders aggregation with the lowest slack and respecting the production constraints have priority for release.

In Figure 2, the backlog of a work centre is represented by an area proportional to the sum of the work hours required by the orders in queue at that centre: each centre is represented by its work load, derived from the customer orders aggregations already released (thus becoming production orders in each centre), plus the load due to the customer orders aggregation whose release has been simulated: the latter load is placed high as it is hypothesized that the work load stack empties from the lower part (in fact, when there is no urgency work is carried out first on the oldest orders).

As the start of customer order processing in a work centre (that is, of its relative production orders) is dependent on the completion of semi-processed pieces worked for that customer order in the work centre immediately above in the line, or on the availability of materials in the intermediary storehouses, it appears that the only admissible situation is that of rising steps of loads downstream in the production system.

Problems arise when work does not take this ideal form, that is when the centres down the line do not receive the materials to be processed in time: if $W(N)$ is the centre N work load (in a defined time unit), N-1 is the centre immediately preceding the centre N and $W(N-1) > W(N)$, when time $W(N)$ has passed without further releases having taken place, the centre N is left without work and must wait for the deposit of the lots still being processed or queuing in the centre N-1. Thus it is necessary to arrange for one production order to start following the completion of the production order (relative to the same customer order) in the immediately preceding

centre. The solution is to raise the work load in the lower centres so as to create a rising steps profile. In the example, the work load of the centre N must be raised by the quantity $W(N-1) - W(N)$.

The supplementary loads (one is shown by dark shading in Figure 2) needed to create the above mentioned profile are not the consequence of requirements derived from a specific order and so must be carefully managed. They must be greater than or equal to the minimum lot size, so line ** (Figure 2) can be only at the same level of line * or higher if the supplementary load needed to create the step is lower than the minimum lot size. These supplementary loads are constituted by free units similar to those already mentioned which were needed to reach the minimum lot size.

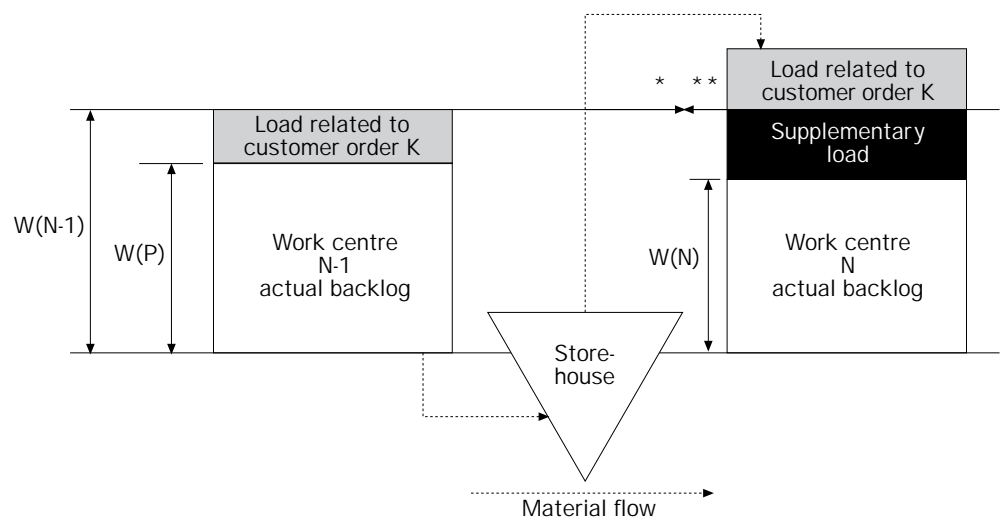
In addition, having fixed the work time for the supplementary load, what type of work should be carried out? There exists a rule which determines the work according to the semi-finished product in the downstream storehouse which differs the most from its usual average.

When a rising step profile has been assured, the slack time for each customer orders aggregation can be easily calculated. The aggregation with the lowest slack have priority for release.

However, it may not be the order with the lowest slack that is first released, as other factors also influence this decision, and these are taken into account by a series of rules, such as global considerations of the client or constraints already described (i.e. completion of all the customer orders within their due-date, processing only one production order at

Figure 2

The working backlog of the centres N and N-1 and the need for a supplementary load



a time on each machine, etc.). There are rules which prevent release if the free units of the supplementary loads needed are higher than a certain global value; as a consequence a certain delay in the completion time can be allowed.

After this analysis, resource allocation is arranged and order release passes from simulation mode to actual mode.

When several production lines converge on one or more final assembly centres, the situation becomes complicated. The slack time is calculated as the difference between the due-date and the work load time of the final centre with the highest work load. The line with the greatest terminal load is thus the slowest. As a consequence, there are two alternatives:

- 1 early loading of the work centres of the faster lines;
- 2 late loading of the work centres of the faster lines.

In the first case, the faster lines are loaded while waiting for the slowest line and when they have performed their jobs deposit the items in the storehouses. In the second case, all the components are deposited at the same instant, even those of the slowest line, so the faster lines are free until the slowest line can work the components relative to the same customer order; during this time, the faster lines are loaded with supplementary loads.

Dynamic rescheduling

A dynamic rescheduling mechanism for the management of urgent orders is provided. Managing a customer order (or a customer orders aggregation) defined as urgent and thus with processing priority over the others,

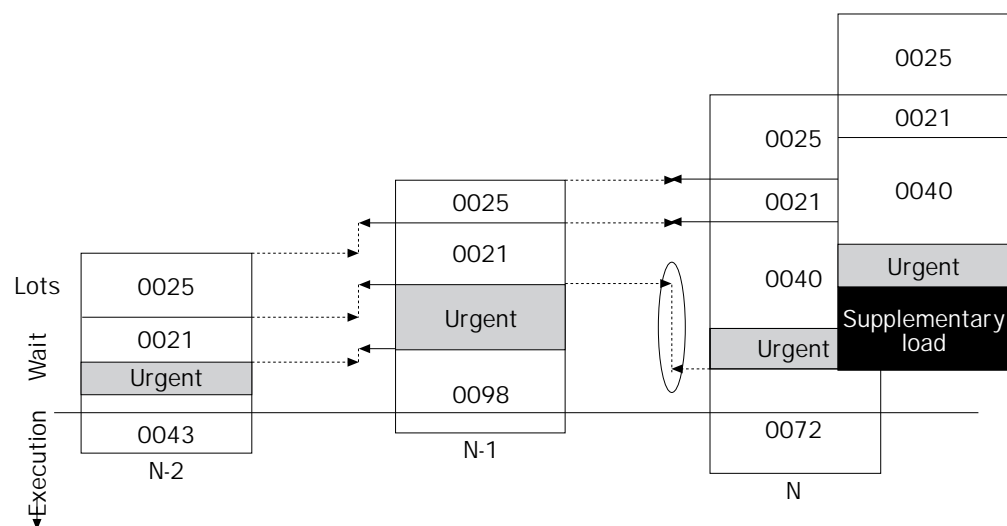
even if they have already been released, poses the problem of modifying all the previous scheduling. Since the production orders, derived from an urgent customer order, must be processed at once, each of them is inserted into the graphs (representing the work load of the centres) immediately above the lots at present being processed, which are one per centre (in the situation illustrated in Figure 3, the numbers of the production orders refer to the respective customer orders aggregation).

This insertion could alter the rising steps relative to every order: in Figure 3 a situation is shown where it becomes necessary to make use of a supplementary load, since, after the insertion of the urgent customer order, the work centre N must start with the urgent customer order, but the corresponding lot (i.e. the production order referring to the same urgent customer order) has not yet been completed in the N-1 work centre. It can also be seen, in Figure 3, how the scheduling of lots referring to other customers' orders (for example 0021 and 0025) proceeds so as the depositing data, for each production order in a work centre, are identical or successive to the start of processing – for the production orders relative to the respective customer order – in the following work centre, even after the consideration of the urgent customer order. In this way, in spite of the complexity of managing lots of different sizes, one has a dynamic rescheduling which ensures the correct procedure of the work-in-process (WIP).

Because too large supplementary loads imply the processing of materials which do not meet actual requirements, it is preferable

Figure 3

The rescheduling mechanism: the insertion of an urgent customer order and the control of the rising steps profile



to keep them as small as possible: in this case the priority for urgent customer orders cannot be accepted, and a rule tests for releases immediately after the lot currently being processed but before other production orders already released.

In summary, supplementary loads can be created for three reasons:

- 1 the customer order aggregation (i.e. minimum lot size needed);
- 2 the production order scheduling (i.e. rising step profile needed);
- 3 the dynamic rescheduling (i.e. urgent orders management).

To prevent there being too many free units in the system, some constraints can be relaxed (i.e. a certain delay in the completion time can be allowed) or some urgencies cannot be accepted.

Discussion

The lot production scheduler proposed can be considered innovative for two reasons:

- 1 it utilizes a hybrid push/pull approach to lot production scheduling, verifying the finite capacity loading and the availability of the materials;
- 2 it exploits the expert system technology in order to obtain satisfactory solutions.

Now, we will examine these facts.

The hybrid push/pull approach

Push systems (e.g. despatching) are those where work is launched in anticipation of a need; pull systems (e.g. a *kanban* system) are those where work is drawn along the production line by downstream consumption i.e. it is downstream usage which triggers movement[1]. In our opinion and according to the meaning of the terms, the main difference between the two systems is how the launches are made, even if some authors[2,3] argue about the different interpretation of a vacancy at a work centre (signal of available capacity in a push system versus signal of usage in a pull system).

Combinations of push and pull strategies in production scheduling are more and more often proposed in the literature, seem to be easier to implement and may achieve better results than either pure push or pure pull strategies. Much of recent research has focused on two major topics:

- 1 how to compare the performances of push and pull systems[4-7]; and
- 2 how to mix the push and pull strategies to achieve advantages[8,9].

Pull systems seem to produce superior results when they can be applied[10]. We can model pull and push systems as a closed and an open queuing network respectively; these models are appropriate since push systems schedule throughput and measure WIP, while pull systems set the WIP levels and measure throughput: it has been demonstrated[11] that the control of WIP is more effective than the control of throughput. In addition, in plants operating near the maximum capacity, throughput remains nearly constant, while WIP – if unchecked – can grow to dangerously high levels.

Unfortunately, pull systems are not applicable to many production environments: they are intrinsically systems for repetitive manufacturing; they will not work in a shop controlled by job orders. For this reason, many applications of the older, and arguably less effective, despatching approach remain or, alternatively, hybrid push/pull strategies have been developed.

For example, a synchro-MRP system, where work is scheduled by despatching but cannot be started without a *kanban* as authorization as well, is described by Hall[12].

The typical hybrid solution[13-15], valid for a general multi-stage production and inventory system, is to use a push strategy at the initial upstream stages and a pull strategy at the downstream stages. Other solutions have been proposed by Goldratt and Fox[16], with the model drum-buffer-rope (DBR), which resembles the proprietary software called OPT [17], and by Spearman *et al.*[18], with the CONWIP.

Under DBR, a drumbeat for the rest of the plant is maintained by sequencing work to be done at the bottleneck station; the drumbeat is then protected by maintaining a time buffer for parts going to the bottleneck; non-bottleneck operations are then scheduled to maintain this buffer (it may be interpreted as a pull strategy). Finally a rope is tied from the bottleneck to material release points to ensure that material is released only at the rate that it is used by the bottleneck (it may be interpreted as a push strategy), thereby preventing an increase in inventory.

While in a *kanban* system each card is used to signal production of a specific part, with CONWIP (CONstant WIP) production cards are assigned to the production line and are not part number cards: part numbers are assigned to the cards at the beginning of the production line by referencing a backlog list; when work is needed for the first work centre in the production line, the card is removed from the queue and marked with the first part number in the backlog for which raw materials are present. Maintenance of the backlog is

the responsibility of production and inventory control staff and in many cases it is generated from a master production schedule (MPS). The queue discipline used at all work centres in the line is the first-in-first-served one, and jobs are pushed between workstations in series once they have been authorized by a card to start at the beginning of the line.

The hybrid model proposed, though part of this line of research, can be considered original because it has a pull management (that is, with releases called for from downstream) along the whole line, while the push part only regards the release of supplementary loads (i.e. the free units) in order to obtain the minimum lot size or the rising steps profile of the work centre loads.

The expert system solution

Several outlines regarding expert system applied to production scheduling have been made[19,20]; some complete schedulers have been constructed (ISIS is probably the most famous expert system for scheduling) and in the literature there are detailed comparisons[21]. A decision support system (DSS) solution has been proposed too[22]. At other times the simple weakness of traditional approaches has been remarked on[23,24]. Particular attention is dedicated to the possibility of an effective rescheduling (or dynamic scheduling)[25–28].

The advantages of the expert system technology in scheduling have been noted by several researchers[29,30] and can be summarized as the possibility of a selective relaxation of the constraints and the use of heuristics to restrict the number of alternatives and assist in selecting the best solution. Resorting to sub-optimal solutions (typical of the expert systems) is necessary since little advance has been made towards finding optimal solution procedures for models of a realistic size. Carrying out research on sub-optimal solutions using heuristic rules would lead to very interesting results[31].

This scheduler is original in the application of a hybrid pull/push approach (rising step profile + supplementary load) by an expert system. The scheduler has a constraint-directed chaining (according to the five classes described by Kusiak[32]: hierarchical, non-hierarchical, script-based, opportunistic and constraint-directed); constraints provide guidance and bounds in the search for good schedules.

This scheduler uses some blackboard techniques of the type hypothesized by Hayes-Roth[33]. The production scheduling blackboard consists of frames, lists and rules of the IF/THEN type, plus a blackboard controller with a shopfloor control system interface and

codes/routeings archives. The status of the work centres, the backlogs and the functional parameters are described by frames; the status of the storehouses and of the order portfolio are instead in the form of lists. The reasoning logic is described by about 200 rules, in lisp-like language. For production scheduling problems, the main advantage of a blackboard control is simplicity of rule drafting and their insertion into the knowledge base, without having to be placed at a precise point in the knowledge base and solely as an enrichment of the knowledge base itself, on which the inferential engine acts.

Findings

We tested the proposed lot production scheduler in the productive context shown in Figure 1, with a large variety of orders incoming frequently requiring about 50 different operations. Experimental conditions are illustrated in the Appendix; major results are shown in Table I. The results were compared with a traditional despatching system, working at finite capacity[34–36]. (Resort to techniques of the *kanban* type were not considered as these have been found to be applied more effectively to production with an elevated degree of repetitiveness, and this was not our case.)

The traditional despatching system works in this way: for each customer order, the lead time is calculated as the sum of the production lead times in all the work centres (composed of the standard time for run, set-up and the average queue time). The slack time (due date minus total lead time) priority rule is applied to despatch the orders. Aggregation of orders is possible only by a human expert, after the despatching.

With the presented scheduler, on the other hand, scheduling takes place after having, from time to time, checked the queue at each centre, and thus it is not based on information concerning the usual average but simulating order release and its real waiting times. Order aggregation is also considered by rules of the expert system.

Table I

The main performances of the two scheduling systems: the pre-existing despatching (DES) and the presented expert lot scheduler (ELS)

	DES	ELS
Orders early	18%	7%
Orders late	32%	12%
Delivery date standard deviation (hours)	4.65	2.66
Average throughput time (hours)	15.7	16.2
Average WIP (× £1,000)	18.3	15.9

The comparison between the proposed scheduler and the pre-existing despatching system was carried out by testing the systems alternatively for a week over an entire period of two months.

The percentage of production orders which did not respect scheduling (both early and late) was found to be about three times lower than that with traditional despatching, although both the average throughput time (order delivery date minus receipt date) and WIP (sum of the storehouse values) remained substantially the same. The standard deviation of the delivery dates was significantly different: with the proposed scheduler it is reduced by almost half.

Therefore our scheduler does not enable one to obtain lower throughput times or a significantly lower WIP, but seems to be considerably better in terms of delivery reliability.

On the other hand, the proposed scheduler has the disadvantage of the required computing time (several minutes for each order release) and the complete monitoring of storehouses and work centre backlogs (which can be very expensive).

Conclusions

The proposed lot production scheduler, based on a model whose implementation and functionality are made possible by the rule-based expert system technology, presents a new way of solving the problems of scheduling lot production. In particular, while the MRP (material requirements planning) system calculates the material requirements independently of the availability of capacity tested by the CRP (capacity requirements planning) system, and with further modifications made by the priority rules of the detail scheduling of the SFC (shop floor control) – despatching, the proposed model could be an interesting alternative, which considers simultaneously the material requirements, the capacity and the scheduling of lots. It does this by means of simulating the queuing times and considering the operational constraints in the search for solutions. It is precisely the presence of queues as a dynamic fact which leads to an intelligent approach. Unlike a despatching system, the presented lot production scheduler simultaneously takes into account both material and capacity in attempting to find a feasible plan: lot sizing and sequencing are done concurrently, though that requires much information and computational effort.

References

- 1 De Toni, A., Caputo, M. and Vinelli, A., "Production management techniques: push-pull

- classification and application conditions", *International Journal of Operations & Production Management*, Vol. 8 No. 2, 1988, pp. 35-51.
- 2 Sarker, B. and Fitzsimmons, J., "The performance of push and pull systems: a simulation and comparative study", *International Journal of Production Research*, Vol. 27, 1989, pp. 1715-31.
- 3 Baker, K.R., Powell, S.G. and Pyke, D.F., "The performance of push and pull systems: a corrected analysis", *International Journal of Production Research*, Vol. 28 No. 9, 1990, pp. 1731-6.
- 4 Lee, L.C., "A comparative study of the push and pull production systems", *International Journal of Operations & Production Management*, Vol. 9, 1989, pp. 5-18.
- 5 Plenert, G. and Best, T.D., "MRP, JIT and OPT: what's 'best'?", *Production & Inventory Management*, Vol. 27, 1986, pp. 22-9.
- 6 Rice, J.W. and Yoshikawa, T., "A comparison of kanban and MRP concepts for the control of repetitive manufacturing systems", *Production & Inventory Management*, Vol. 23, 1982, pp. 1-13.
- 7 Sillince, J.A.A. and Sykes, G.M.H., "Integrating MRPII and JIT: a management rather than technical challenge", *International Journal of Operations & Production Management*, Vol. 13 No. 4, 1993, pp. 18-31.
- 8 Discenza, R. and MacFadden, F.R., "The integration of MRPII and JIT through software unification", *Production & Inventory Management*, Vol. 29, 1988, pp. 49-53.
- 9 Gelders, L.F. and Van Wassenhove, L.N., "Capacity planning in MRP, JIT and OPT: a critique", *Engineering Costs and Production Economics*, Vol. 9, 1985, pp. 201-09.
- 10 Krajewski, L.J., King, B.E., Ritzman, L.P. and Wong, D.S., "Kanban, MRP, and shaping the manufacturing environment", *Management Science*, Vol. 33, 1987, pp. 39-57.
- 11 Karmarkar, U.S., "Lot sizes, lead times and in-process inventories", *Management Science*, Vol. 33, 1987, pp. 409-23.
- 12 Hall, W.R., *Zero Inventories*, Dow Jones-Irwin, Homewood, IL, 1983.
- 13 Hodgson, T.J. and Wang, D., "Optimal hybrid push/pull control strategies for a parallel multistage system: Part I and Part II", *International Journal of Production Research*, Vol. 29 Nos 6/7, 1991, pp. 1279-87 and pp. 1453-60.
- 14 Hirakawa, Y., Hoshino, K. and Katayama, H., "A hybrid push/pull production control system for multistage manufacturing processes", *International Journal of Operations & Production Management*, Vol. 12 No. 4, 1992, pp. 69-81.
- 15 Lee, C.Y., "A recent development of the integrated manufacturing system: a hybrid of MRP and JIT", *International Journal of Operations & Production Management*, Vol. 13 No. 4, 1993, pp. 3-17.
- 16 Goldratt, E.M. and Fox, R.E., *The Race*, North River Press, New York, NY, 1986.
- 17 Goldratt, E.M., "Optimized production timetable (OPT): a revolutionary program for

- industry", *APICS Annual Conference Proceedings*, 1980, pp. 172-6.
- 18 Spearman, M.L., Woodruff, D.L. and Hopp, W.J., "CONWIP: a pull alternative to kanban", *International Journal of Production Research*, Vol. 28 No. 5, 1990, pp. 879-94.
 - 19 Biggs, J.R., "Heuristic lot sizing and sequencing rules in a multi-stage production-inventory system", *Decision Science*, Vol. 10 No. 1, 1979, pp. 96-115.
 - 20 Fiedler, K., Galletly, J.E. and Bicheno, J., "Expert advice for JIT implementation", *International Journal of Operations & Production Management*, Vol. 13 No. 6, 1993, pp. 23-30.
 - 21 Kathawala, Y. and Allen, W.R., "Expert systems and job shop scheduling", *International Journal of Operations & Production Management*, Vol. 13 No. 2, 1993, pp. 23-35.
 - 22 Hendry, L. and Kingsman, B., "A decision support system for job release in make-to-order companies", *International Journal of Operations & Production Management*, Vol. 11 No. 6, 1991, pp. 6-16.
 - 23 Melnyk, S.A. and Piper, C.J., "Lead time errors in MRP: the lot sizing effect", *International Journal of Production Research*, Vol. 23 No. 2, 1985, pp. 253-64.
 - 24 Melnyk, S.A. and Ragatz, G.L., "Order review/release and its impact on the shop floor", *Production & Inventory Management*, Vol. 29, 1988, pp. 13-17.
 - 25 Brown, M.C., "The dynamic reproduction scheduler: conquering the changing production environment", in Pau, L.F., Motiwalla, J., Pao, Y.H. and Theh, H.H. (Eds), *Expert Systems in Economics, Banking and Management*, North-Holland, Amsterdam, 1989.
 - 26 Huang, Y.G. and Kanal, L.N., "Dynamic scheduling problem solving: an object-oriented approach", in Balagurusamy, E. and Howe, J.A.M. (Eds), *Expert Systems for Management and Engineering*, Ellis Horwood Limited, Chichester, 1990.
 - 27 Sarin, S.C. and Salgame, R., "A knowledge-based system approach to dynamic scheduling", in Kusiak, A. (Ed.), *Knowledge-based Systems in Manufacturing*, Taylor & Francis, Philadelphia, PA, 1989.
 - 28 Szelke, E. and Kerr, R.M., "Knowledge-based reactive scheduling", *Production Planning & Control*, Vol. 5 No. 2, 1994, pp. 124-45.
 - 29 Steffen, M.S., "A survey of artificial intelligence-based scheduling systems", Fall Industrial Engineering Conference, Boston, MA, 7-10 December 1986.
 - 30 Meyer, W. (Ed.), *Expert Systems in Factory Management - Knowledge-based CIM*, Ellis Horwood, Chichester, 1990.
 - 31 Zeestraten, M.J., "The look ahead dispatching procedure", *International Journal of Production Research*, Vol. 28 No. 2, 1990, pp. 369-84.
 - 32 Kusiak, A. (Ed.), *Knowledge-based Systems in Manufacturing*, Taylor & Francis, Philadelphia, PA, 1989.
 - 33 Hayes-Roth, B., "A blackboard architecture for control", *Artificial Intelligence*, Vol. 26, 1985, pp. 251-321.
 - 34 Blackstone, J.H. Jr, Phillips, D.T. and Hogg, G.L., "A state-of-the-art survey of dispatching rules for manufacturing job shop operations", *International Journal of Production Research*, Vol. 20 No. 1, 1982, pp. 27-45.
 - 35 Gardiner, S.C. and Blackstone, J.H. Jr, "The effects of lot sizing and dispatching on customer service in an MRP environment", *Journal of Operations Management*, Vol. 14 No. 11, 1993, pp. 143-59.
 - 36 Kanet, J.J. and Hayya J.C., "Priority dispatching with operation due dates in a job shop", *Journal of Operations Management*, Vol. 2 No. 4, 1981, pp. 167-75.

Appendix Experimental conditions

Nine work centres, two final assembly centres, three productive lines.
Different working parameters for each centre, independent of the type of the operation performed:
minimum lot size, set-up time, run time per unit;
53 different operations;
142 different end-items sold.
Input data (for each centre)

minimum lot size	both DES and ELS
set-up time	both DES and ELS
current set-up	ELS
run time per unit	both DES and ELS
average queue time	DES
downstream storehouse level	both DES and ELS
work load (backlog)	ELS
free units	
in the downstream storehouse	ELS
in the work load (backlog)	ELS

Input data (for each order)

receipt date	
type	
quantity	
due-date	
delivery date	

Table A1
Weekly scheduling

Week	Scheduling system adopted	Number of orders received	Number of different end-items
1	DES	225	68
2	ELS	246	79
3	DES	180	123
4	ELS	168	106
5	DES	193	84
6	ELS	141	90
7	DES	205	102
8	ELS	174	95

(DES = pre-existing despatching; ELS = proposed expert lot scheduler)