Design of maintenance control systems

Department of Operations Planning and Control -- Working Paper series

Design of Maintenance Control Systems

C.W. Gits

Research Report TUE/BDK/LBS/93-31

Graduate School of Industrial Engineering and Management Science
Eindhoven University of Technology
P.O. Box 513, Paviljoen F16
NL-5600 MB Eindhoven
The Netherlands

Summary

Maintenance control deals with the coordination of maintenance demand and resources in such a way that stated objectives are satisfied. This allocation problem is characterized by complexity of the objectives, uncertainty of demand and flexibility of the capacity resources. These conditions warrant a hierarchical decision structure primarily aimed at a stepwise refinement of control on the basis of increasingly detailed information. A reference framework of decision functions is developed to facilitate the design of maintenance control systems.

1. Introduction

Maintenance in an industrial organization is essentially concerned with the availability of equipment for production purposes. In many organizations the production process has changed drastically in recent years. Attention shifted from increasing efficiency by means of economies of scale and internal specialization to meeting the market conditions in terms of flexibility, delivery performance and quality. This trend implies working with less and less inventories. Consequently, unavailability of means of production will result in serious delivery problems more often. One way of dealing with unavailability is getting rid of it as much as possible. This can be achieved by modification of the system eliminating underlying failures. Another way is to minimize the consequences of unavailability by prescription of preventive maintenance transforming unplanned into planned unavailability. However, both approaches can not eliminate unplanned unavailability altogether. Production and maintenance control have to explicitly anticipate its occurrence by creating flexibility.

In general, existing maintenance control systems are not geared to achieving flexibility. Their focus still is on realizing economic objectives. This implies planning as much work as possible in an attempt to realize the high utilization rate of the maintenance capacities so highly favored by management. A high utilization rate, however, either results in urgent work having long and unreliable lead times or planned activities being rescheduled time and time again. In any case, the controllability of the total amount of work is insufficient to meet the stricter requirements of production. Often it is attempted to solve this problem by means of improving the information processing capabilities. However, the results of this approach are rather disappointing.

Production control which has been studied more comprehensively is faced with an analogous allocation problem. Here, too, solutions were proposed based on improving the supply of information (Orlicky [1], Plossl/Welch [2]). The problems encountered in implementing these concepts in real situations strongly indicate the need to review critically the structure of production control (Meal [3], Bertrand/Wortmann/Wijngaard [4], [5]).

This paper aims at establishing a reference framework of decision functions to guide the design of maintenance control systems suited to meet the altered requirements of production. It will draw heavily upon knowledge acquired in production control. In section 2 the maintenance process is characterized and the basic terms are defined. The main elements of maintenance control i.e. demand, resources and objectives are discussed in section 3. The reference control structure consists of six decisions which are presented in section 4. Finally, in section 5 conclusions are drawn and proposals for further study formulated.

2. Maintenance

Maintenance in an industrial organization (fig. 1) has the objective of supporting the production process in which the primary production input (material, energy, manpower) is transformed into the primary production output (the desired product). This transformation process makes use of a diversity of technical production systems. A *technical production system* is a collection of physical elements with a specific production function. The *condition* of a technical production system is the physical ability considered relevant for fulfillment of its function. External causes, ageing and use impair the condition of the technical production systems, inevitably resulting in failures of the system. *Failure* denotes the transition of a technical production system to the condition in which it is inadequate for fulfillment of its function. Of course, it may be possible to shutdown a system before the occurrence of failure. Either way, production leads to a secondary output: demand for

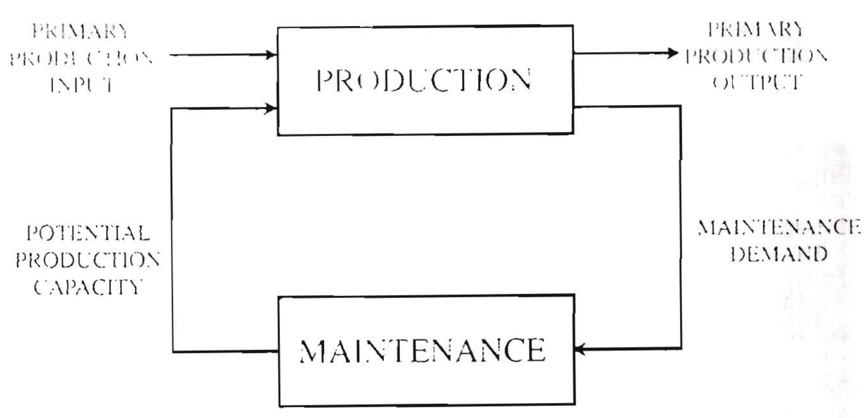

Fig. 1: The relationship between production and maintenance

maintenance. On being carried out by the resources of maintenance, this leads to a secondary production input; that is potential production capacity. In this view, *maintenance* is the total of activities aiming to retain the technical production systems in or restore them to the condition necessary for fulfillment of the production function. In this definition, "retaining in" corresponds to preventive maintenance, "restoring to" with corrective maintenance.

3. Maintenance control

Maintenance control (fig. 2) essentially deals with the operational coordination of demand and resources to achieve stated objectives.

Of course higher levels of decision making can be found in an organization such as strategic planning and management control (Anthony [6]). Strategic planning is primarily concerned with the general direction in which the organization as a whole is heading. Management control deals with the integration of the various processes within the organization. However, these decisions do not directly influence maintenance execution. Their impact is in the way they shape the environment of maintenance control i.e. demand, resources and objectives.

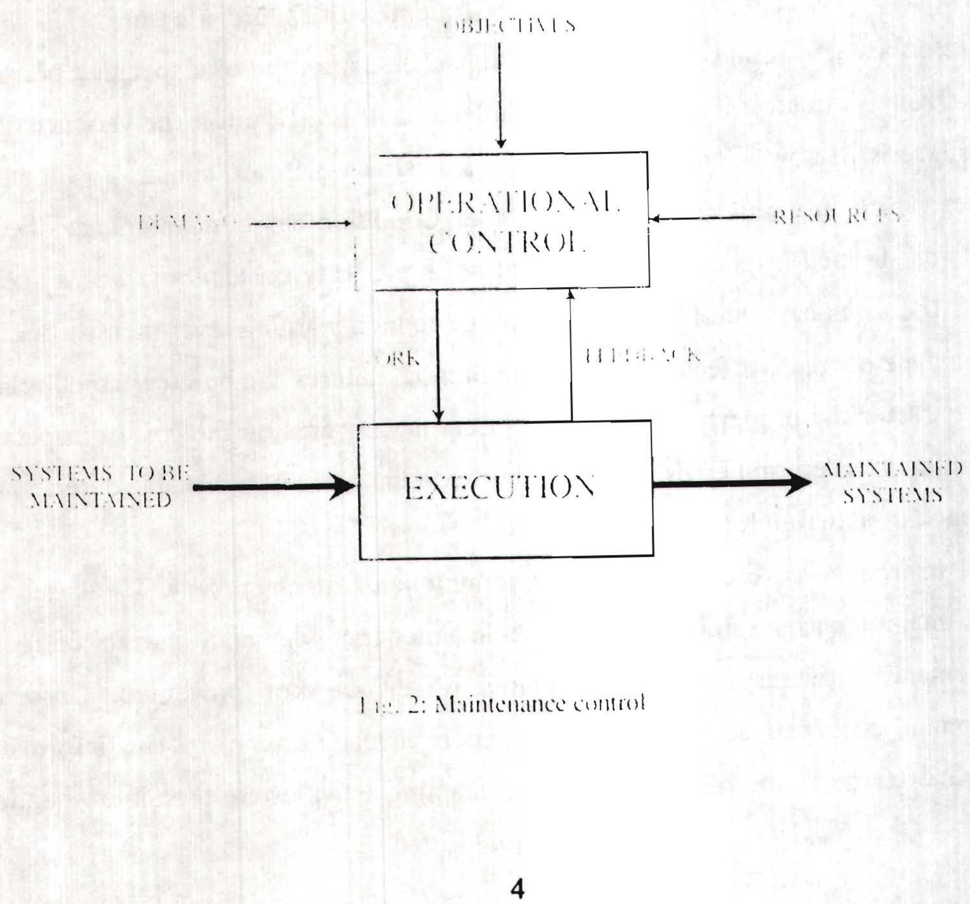

Fig. 2: Maintenance control

3.1. Demand

Maintenance demand specifies what operations should be carried out and when. It is the total demand for the individual technical production systems in the organization. These systems are a closed group which is known beforehand. Maintenance demand is the sum of the demand of the individual systems. Information on the future demand of an individual technical production system is to be had from its operation intensity and from its maintenance concept. The operation intensity follows from the production plan and is autonomous. The maintenance concept is the set of rules on what operations are to be carried out and how demand for these operations is activated (Gits [7]). The program of requirements to be satisfied by each concept is specified at higher levels of control. From the point of view of maintenance control, demand has to be regarded as given.

With respect to the eventual demand for resources, three types of maintenance rules have to be distinguished:

- failure-based maintenance;
- use-based maintenance;
- condition-based maintenance.

Failure-based maintenance prescribes activation of an operation in the event of failure. This type of rule is always effective. It is efficient if the consequences of failure are small. Use-based maintenance prescribes activation of a well defined repair on expiration of a specified period of use. This type of rule is effective if the failure rate increases. It is efficient if the variance about mean use to failure is narrow. Generally, it is concerns parts which are subject to fatigue or mechanical wear which are well correlated with use. Condition-based maintenance dictates activation of a well defined inspection of a characteristic property on expiration of a specific period of use. If the assessed value has passed a predetermined warning level then recondition is necessary. This type of rule is effective only when potential failures can be ascertained reliably by means of a characteristic property. Its efficiency primarily concerns on the cost of inspection. The inspection interval is determined by the period between the onset of noticeable deterioration and the occurrence of actual failure.

Failure-based maintenance results in corrective maintenance demand only. Use-based and condition-based maintenance result in preventive maintenance demand and in corrective maintenance demand with respect to the failures which are not prevented. Corrective maintenance demand occurs unexpectedly, its contents is virtually unknown. This information becomes available after execution of diagnosis; The remaining activities are then known and the resources needed can be determined with some accuracy. In principle, the moment of preventive maintenance demand is predictable. This also holds for its contents. However, it should be noted that in the case of condition-based maintenance this information is not complete until after execution of inspection. Only then it is known whether or not recondition has to be carried out.

3.2. Resources

Maintenance resources consist of materials and capacities. Material resources are items which are consumed in the execution of maintenance operations and have to replenished. Capacity resources are the personal abilities, instruments and facilities which are used during execution of maintenance operations and which can be used time and time again. In reality, the distinction between materials and capacities is not as clear cut as implied here. Some materials ("recoverables") can be maintained and used more than once; Capacities are depleted in the long run and have to be renewed. However, decisions with respect to these aspects are considered to be outside the scope of this paper.

Two categories of materials may be ordered on the basis of planned or actual demand: specials and adequate warning items. Specials are items which have been bought for use on a specific date, e.g. in preparation for a major overhaul. Adequate warning items are items which either have a minor failure but can be economically patched up for a period longer than the ordering lead-time, or their wear indicates, by a period longer than the lead-time, impending failure. The ordering lead-time of the remaining items is too long to allow for coordination with demand, actual or planned. These items have to be stocked. The decisions how many items to stock and when to order have to be based on forecasts of demand alone. Coordination of material replenishment and demand is based on static characteristics. At the operational level, information about demand for items is passed on to the materials management function to be fulfilled. If demand cannot be met, the technical production system concerned remains unavailable until the item is purchased. The consequences of stock-out may be minimized by special activities such as installing the next higher assembly, cannibalization of another technical production system, substitution with a higher grade item. However, these activities do not yield flexibility as they do not anticipate the problem at hand.

Structurally, the amount of capacities to be employed in meeting demand is dependent on factors such as illness, training, holidays. The conflict between effective and efficient use of the capacities is solved by specifying *operational constraints*. These constraints concern aspects such as utilization level, work-in-progress, capacity variations. A high level of utilization of these capacities will reflect favorably on maintenance economics. However, it also results in long and unreliable maintenance throughput times frustrating production. In order to achieve the norm rate decided upon, it is necessary to allow a certain amount of work inventory. This working stock creates lead times which are longer than strictly necessary. As the inflow of work can only be influenced up to a certain degree and the actual capacity needs become known at a very late stage, adjustment of capacities by making use of their flexibility is a necessity. Volume-flexibility implies that during periods the amount of a type of capacity can be increased by means of contracting-in personnel or by working overtime. Each realization may be possible to a limited extend, at specific cost and with a specific preparation time. Mix-flexibility exists if the maintenance engineers are multi-functional i.e. master a number of skills. The amount of capacity of a specific type can be increased simultaneously decreasing the amount of another type. Mix-flexibility can be used in the very short term but may result in work taking longer to be completed.

3.3. Objectives

The objectives of maintenance control have to be derived from production, the primary process in the organization. The objectives of production are twofold. On the one hand, the production-system should show high flexibility with respect to aspects such as changing market conditions, fluctuating production demand forecasts, actual demand variations. Lack of flexibility may lead to high and unbalanced stocks, poor delivery performance and possibly loss of market. On the other hand, realizing acceptable production costs is a sine qua non for survival of the organization in the long run.

The contribution of maintenance control to high flexibility of the production-system focusses on realizing high availability of the technical production systems for manufacturing purposes. The weigh attributed by production to the availability of a technical production system depends on its essentiality for the production process. Maintenance of the technical production systems considered to be of the utmost importance has to be carried out according to requirements of production laid down at higher levels of control. These requirements limit preventive maintenance execution to planned non-productive periods. Corrective maintenance has to be carried out as soon as possible. The coordination of production and maintenance of the less essential technical production systems focusses on the start and duration of maintenance execution. Preventive maintenance has to be started before the date laid down in the demand. The start of corrective maintenance execution can be delayed for some time without problems . Conditions with respect to the duration of maintenance are not too sharp and primarily depend on the technical characteristics of the operations to be executed. The eventual timing depends not only the wishes of production but also on the possibilities of maintenance. How these wishes and possibilities are to be reconciled has been decided upon at higher levels of control.

Given the operational constraints, the efficiency of maintenance control is essentially concerned with the allocation of capacity to maintenance flows and operations. The amount of capacity which can be freely allocated to preventive maintenance depends on the amount of corrective maintenance expected to turn up eventually and the average response-time specified at a higher level of control. Furthermore, depending on the agreements governing the relationship between production and corrective maintenance, additional norms may have to be set with respect to the distribution of its load and tightness. The allocation of capacity to individual operations is faced with a high uncertainty as far as failure-based and condition-based maintenance are concerned. The uncertainty about the task to be carried out implies that the capacity needs can be estimated up to a certain degree only. A reduction of this uncertainty can be achieved by treating diagnosis/inspection and repair/recondition as two distinct activities each requiring capacity. The outcome of diagnosis/inspection forms the basis for the allocation of capacity to repair/recondition. The benefits in terms of improved capacity use have to be traded off against the losses in terms of increased control effort and, possibly, additional set-ups. A refinement is the allocation of a certain capacity budget. If diagnosis/inspection shows that the total operation requires more capacity then feedback is required and additional capacity can be granted or not depending on the actual state of affairs.

4. Maintenance control system

A maintenance control system encompasses all decisions with respect to the timely coordination of demand and resources taking into account the structural requirements of production and the operational constraints in view of capacities. Such a system should be able to cope with the intricacies of the allocation problem at hand: complexity, uncertainty and flexibility. The complexity results from the effect of the decisions on the objectives being ambiguous and interdependent. Uncertainty about the timing and contents of demand is induced by corrective maintenance and by condition-based maintenance and lingers even during execution. Flexibility of capacity may be available with a variety of lead times and prices. These characteristics favor putting decision power at the lower levels in the organization to be able to react directly to new and more detailed information becoming available. A hierarchical control system consists of a mix of the following four elements (Galbraith [8]):

- creation of self-contained tasks;
- creation of slack resources;
- investment in vertical information systems;
- creation of lateral relations.

Self-contained tasks are tasks which are carried out irrespective of other tasks. Slack resources enhance the decision freedom and can be introduced by means of overcapacity, safety times etc. Vertical information systems concern decision support systems geared to specific functions. Lateral relations, finally, concentrate on coordination of the distinguished systems and the progress of work.

The maintenance control problem will be decomposed into a number of hierarchically ordered decision functions together with the creation of slack necessary to make these functions sufficiently self-contained. Vertical information systems and lateral relations increasing the information processing capabilities have to be developed within this structure. These elements however are not treated in this paper.

The reference control structure consists of the following six decision functions (fig. 3):

- aggregate maintenance control
- preventive maintenance planning;
- corrective maintenance classification;
- work order release;
- work order scheduling;
- work dispatching.

4.1. Aggregate maintenance control

Aggregate maintenance control primarily aims at avoiding potential problems with respect to the availability of bottleneck capacities at the lower level decisions. The availability of capacity is safeguarded by generating norms governing its allocation and adjustment.

Due to the unpredictability of corrective maintenance, it is only possible to reserve a constant amount of capacity per period for this flow. The remaining capacity can be allocated freely to preventive maintenance. The amount of capacity needed for corrective and preventive maintenance is to be had from the production plans specifying operation intensities and from maintenance concepts as basis for the estimation of demand. Although dealing with aggregates, it is necessary to consider the essential technical production systems individually due to the strict production requirements. Capacity shortage may call for upward referral or may be solved by

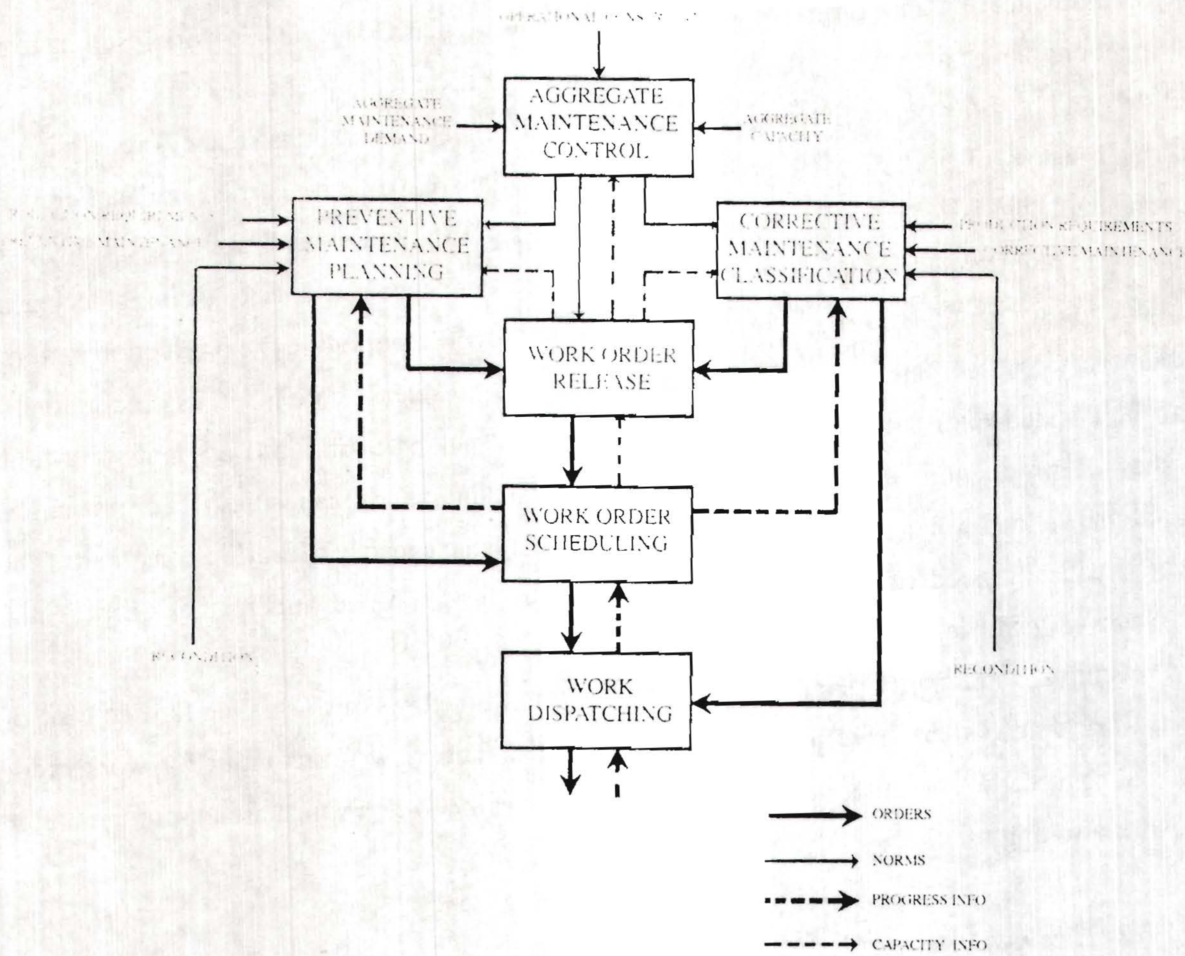

Fig. 3: The structure of maintenance control

making use of capacitative flexibility. The structural relationship between production and corrective maintenance may require setting norms with respect to the distribution of load and tightness of accepted work.

The high degree of uncertainty faced in detailed maintenance control necessitates measurement of its performance. Its performance is reflected in the load and tightness of the work waiting to be released. Feedback of this information to the aggregate level allows adjustment of the capacity, if necessary.

4.2. Preventive maintenance planning

Preventive maintenance planning deals with specifying preventive maintenance work orders. A *work order* essentially specifies what has to be carried out and when. This decision has to see to efficiently meeting maintenance demand, the requirements of production and the capacity allocation constraints.

Maintenance demand consists of work initiated by maintenance concepts, of work resulting from inspections and of overdue work requiring re-planning. Efficiency concentrates on trading off the efforts and benefits of simultaneous execution of maintenance operations. The efforts are in the form of a loss of potential units of use, some operations have to be carried out earlier than demanded. The benefits accrue from a reduction in set-up activities and, possibly, throughput-time. The requirements of production in of view of preventive maintenance of essential technical production systems dictate execution during the non-productive periods in the production pattern. Maintenance of less essential systems is not critical. A latest release date is set, its precise timing can be decided upon later at a lower level of decision making taking into account short-term wishes of production and the actual state of the maintenance capacities. The preventive maintenance load which can be planned is constrained by the capacity allocation constraints. The overall capacity need of preventive maintenance may be smoothed anticipating on potential problems at lower levels of decision making.

Preventive maintenance planning eventually results in two flows of work. One flow consists of *fixed work orders* which have to be carried as planned. This flow forms input for work order scheduling directly. The other flow consists of *advancable work orders* which are input for a buffer of work awaiting release.

4.3. Corrective maintenance classification

Corrective maintenance classification is concerned with generating corrective maintenance work orders. This decision has to reflect efficiently the wishes of production, the capacity allocation constraints and the actual condition of maintenance.

Maintenance demand consists of work initiated by failures, of work resulting from diagnosis and of overdue work requiring re-classification. Failures of an essential technical production system have to be rectified as soon as possible, leaving no play whatsoever. With respect to the less essential systems, production and maintenance have to reach an agreement about the latest start-date of the work involved. Production has to translate the importance of the system for their day-to-day operations into an acceptable duration of unavailability. Maintenance has to consider this wish in the light of the capacity needs of already accepted but not yet realized work and the norms with respect to the load and tightness. If these norms leave no room then production has to lower its wishes. Efficiency of maintenance focusses on the amount of capacity to be allocated to restoring a failure and on whether or not diagnosis and repair are to be considered as one work order.

Corrective maintenance classification results in two distinct flows of work. One flow consists of *rush work orders* which have to be carried out straightaway and as fast as possible. These work orders are input of work dispatching directly. The other flow encompasses *postponable work orders* which are input of a buffer of work awaiting release.

4.4. Work order release

The work order release function periodically sets free work orders from the buffer of accepted orders to compete for the scarce resources. It aims at keeping the released amount of work under control and consequently the timing of work orders.

Each period the amount of work released should be proportional to the capacity expected to be available in the next period. To compensate for the inevitable fluctuations in rush work orders and deviations in fixed work orders, some feedback to aggregate maintenance control should exist. This information about total load and tightness of the work in the buffer may lead to adjustment of capacity levels. The decision which specific work orders from the buffer are going to be released depends on material availability and order priorities. Only orders demanding material that is available are to be considered, of course, as far as demand is known at this stage. The priority of a work order is given by the remaining slack in view of its latest release-date. Unexpected production interruptions may give rise to additional opportunities to release specific orders.

4.5. Work order scheduling

Work order scheduling deals with the detailed timing of each work order. The input of work order scheduling consists of the fixed work orders generated in preventive maintenance planning and the formally released work orders. The timing of fixed work orders has been taken care of at the planning level. In scheduling these orders only the remaining play, if any, can be used. In the scheduling process, short-term production wishes in view of the availability of technical production systems have to be accounted for as well as short-term capacity bottlenecks, environmental conditions and efficiency considerations.

The eventual schedule specifies start and due-dates or the periods in which work orders have to be carried out. It should be stressed that in view of the high degree of uncertainty inherent in maintenance, the aim of work order scheduling is to set forth a basis for work order dispatching. Generally it will not be possible to adhere to the schedule; problems in work progress may require re-scheduling of work orders or if orders are going to be overdue may require upward referral to planning and classification.

4.6. *Work dispatching*

Work dispatching is the control function concerned with the sequencing of the work orders and allocating each order to a specific capacity type.

As long as the execution of work progresses according to the work order schedule, sequencing of work orders can be based on their tightness. Rush orders always go directly to the top of the stack. A work order will be allocated to a specific capacity type if that type becomes available. Fixed and rush work orders may require interrupting work being carried out on non-essential technical production systems ("preemptive priority"). To avoid work orders waiting for a capacity type to become idle or on going work to be interrupted, it can be decided to make use of mix-flexibility, if available of course.

However, if progress deviates from the schedule then it is no longer a sound basis for decision making alone. When work progresses faster than expected, work orders can be advanced or the capacity can be left idle. With progress lagging behind, this information has to be fed back to scheduling. The sequence in which the orders already under way are going to be finished depends on production wishes and efficiency considerations such as the amount of work already completed and set-ups involved.

From the point of view of efficiency, it may be desirable to dispatch a number of work orders simultaneously instead of on a piecemeal basis. At the one hand, the actual state of affairs becomes ambiguous, on the other hand less time is lost in reporting the progress of work and some decision freedom remains enriching the task to be carried out.

5. Conclusion

Keywords in maintenance control are complexity, uncertainty and flexibility. The complexity results from the effects of decisions on the objectives being ambiguous and the interdependence of these decisions. Uncertainty about the timing and contents of demand is inherent in maintenance; it lingers until its execution is finished. Flexibility of the capacities in various forms may be available to counter the inevitable variations in demand. The above-mentioned characteristics favor a hierarchical structure of the maintenance control system. Such a system allows a stepwise refinement of control; Global decisions based on aggregate information restrict detailed decisions at lower levels of control. Such an approach leaves room for flexible responses to deal directly with all kinds of "disturbances". As new information arises continuously, even during execution of maintenance, feedback and progress monitoring are essential ingredients.

It should be stressed that each situation to be controlled is unique and essentially requires a customized solution. However, a number of aspects are considered to be relatively standard making it possible to formulate a reference framework of decision functions. This framework is a tool to support the design of maintenance control systems. The output and mutual interdependence of the distinguished decision functions are determined by the characteristics of the actual situation. The hierarchy in decision functions does not imply a hierarchy in organizational functions; A number of decision functions can be allocated to one and the same decision maker.

To avoid making the discussion even more complicated than it already is, it was implicitly assumed e.g. that no maintenance is contracted-out, the essentiallity of systems is given, failure results in breakdown of the technical production system, maintenance execution requires shutdown of the technical production system, secondary damage is absent, the relationship between demand, work orders and capacity types is simple. Relieving these assumptions does not seriously effect the framework; It does complicate the decisions.

Practical research is being directed at the applicability of (parts of) the framework in designing maintenance control systems. This requires the design of organisable and implementable decision functions taking into account the characteristics of the situation to be controlled (Dekkers [9], Stuursma [10]). Further theoretical research is needed into the higher levels of control specifying the structural requirements and operational constraints and into aggregate maintenance control balancing the flows of maintenance demand.

Literature

[1] Orlicky, J., Materials requirements planning. Mcgraw Hill, New York, 1975

[2] Plossl, G.W., W.E. Welch, The role of top management in the control of inventory. Reston Publ. Co., Reston, 1979.

[3] Meal, H.C., Putting production decisions where they belong. Harvard Business Review, Vol. 84, pp. 102-111, 1984.

[4] Bertrand, J.W.M., J.C. Wortmann and J. Wijngaard, Production control; A structural and design oriented approach. Elsevier, Amsterdam, 1990.

[5] Bertrand, J.W.M., J.C. Wortmann and J. Wijngaard, Production control and materials management, Stenfert Kroese, Leiden, 1990, (in Dutch).

[6] Anthony, R.N., Planning and control systems: A framework for analysis. Harvard University Press, Boston, 1965.

[7] Gits, C.W., Design of maintenance concepts. International Journal of Production Economics, Vol. 24, pp. 217-226, 1992.

[8] Galbraith, J.R., Designing complex organizations. Addison Wesley, Reading. 1973.

[9] Dekkers, R.W.E., Operational maintenance control at Schiphol Airport. M.Sc. thesis, Eindhoven University of Technology, Eindhoven, 1993, (in Dutch).

[10] Stuursma, W.J., Implementing response times at Schiphol Airport. M.Sc. thesis, Eindhoven University of Technology, Eindhoven, 1993, (in Dutch).