Structuring of Maintenance Control Systems

C.W. Gits
ABSTRACT
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, uncertainty and flexibility. 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 in what essentially are unique situations.

KEYWORDS: Maintenance, Hierarchical Control, Production Management.

INTRODUCTION
Maintenance in an industrial organization has to keep the means of production fit so that they can fulfill their function. 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 market conditions in terms of flexibility, delivery performance and quality. Ideally, this trend towards "Just-In-Time" production implies working without inventories at all. Consequently, unplanned unavailability of means of production will directly result in serious delivery problems. One way of dealing with unplanned unavailability is trying to get rid of it as much as possible. This can be achieved by modification of the means eliminating underlying failures and/or by prescription of preventive maintenance transforming unplanned into planned unavailability. Another way is to minimize the consequences of unplanned unavailability. This requires that both production- and maintenance control explicitly anticipate its occurrence. Both approaches require the creation and control of flexibility in order to be able to deal with the unavoidable short term variations on longer term plans.

In general, existing maintenance control systems are not geared to achieving flexibility. The 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 favoured 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 requirements of "Just-in-time" 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 (Meel [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 stricter requirements of production. It will make extensively use of insights gained in production control. To start with, maintenance is defined and the main elements of maintenance control i.e. demand, resources and goals are characterized. The proposed reference decision structure consists of two levels: work order coordination and maintenance unit control. Work order coordination concentrates on the operational tuning in of production and maintenance, maintenance unit control on realizing the agreed performance. Finally conclusions are drawn, assumptions explicitated and proposals for further study formulated.

MAINTENANCE

Maintenance in an industrial organization (fig. 1) supports 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 systems. A technical system is a collection of physical elements with a specific production function which is regarded as an entity from the point of view of maintenance. The state of a technical system is the physical ability considered relevant for fulfilment of its function. External causes, ageing and use impair the state of the technical systems, inevitably leading to a secondary production output: demand for maintenance. On being carried out, 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 systems in or restore them to the state necessary for fulfilment of the production function. In this definition, "retaining in" corresponds to preventive maintenance, "restoring to" with corrective maintenance.
Maintenance as a secondary process has to contribute to achieving the objectives of production which in their turn have to be in line with the goals of the organization as a whole. These goals, together with the strategies for attaining them, are formulated at the highest level of decisionmaking. To carry out these strategies effectively and efficiently, the various processes within the organization have to be integrated. Strategic planning and management control (Anthony [6]) together result in setting the objectives to be realized by each distinguished process, in defining the available resources and in formulating restrictions with respect to the effective and efficient use of these resources. Given the the outcome of the higher levels of decisionmaking, operational control finally refers to the effective and efficient execution of specific activities.

take in fig. 2.

MAINTENANCE CONTROL

Maintenance control (fig. 2) essentially deals with the operational coordination of demand and resources to achieve the stated goals in terms of effectivity and efficiency.

Demand

Maintenance demand specifies what maintenance operations should be carried out and when. It is the total demand for the individual technical systems in the organization. These systems are a closed group which is known beforehand. Consequently, maintenance demand is the sum of the demand of the individual systems. Information on the future maintenance demand of an individual technical system is to be had from its operation intensity and from its maintenance concept. The operation intensity follows from the production plan. The maintenance concept is the set of rules on what maintenance operations are to be carried out and how demand for these operations is activated (Gits [7]).

Three types of maintenance rules can be distinguished:

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

Failure-based maintenance prescribes activation of maintenance in the event of failure. Failure denotes the transition of a technical system to the state in which it is inadequate for fulfillment of its function. 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 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 norm 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.

With respect to maintenance demand, failure-based maintenance results in corrective maintenance only. Use-based and condition-based maintenance result in preventive maintenance and in corrective maintenance 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. The moment of preventive maintenance demand is, in principle, 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 wether or not recondition has to be carried out. Consequently, both failure-based and condition-based maintenance imply uncertainty about the actual contents of maintenance to be carried out even after execution has started.

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 ("rotables") 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 (Mitchell [8]). 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. Consequently, coordination of material replenishment and maintenance demand is based on static characteristics, the dynamics of the day-to-day operations can not be accounted for. Information about demand for items is passed on to the inventory control function to be fulfilled. If demand cannot be met, the technical system concerned remains unavailable for production purposes 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 system, substitution with a higher grade item. However, these activities do not yield flexibility as they do not anticipate the problem at hand.

The capacities are considered to be organized into one maintenance unit which takes care of total demand for maintenance. The amount of capacities to be employed in meeting demand decided upon at management control level can be varied by making use of flexibility in the capacities. Volume-flexibility implies that during periods the amount of a type of capacity can be increased by means of contracting-out work, by 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 multifunctional 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 and consequently it is a convenient means to counter the daily disruptions in the progress of work. The amount of capacitative flexibility up to a high degree determines the complexity of the coordination of capacity and demand. A high degree of flexibility allows to express the potential of the maintenance unit to carry out work in a small number of bottleneck capacities. A bottleneck capacity is characterized by a high utilization and small flexibility. In view of the other capacities it can be assumed that the amount can be adapted according to demand.

Goals

The goals of maintenance consist of realizing its objectives and and at the same time meeting the operational constraints. The objectives have to be derived from the objectives of production being the primary process. On the one hand, the production-system should show high flexibility with respect to aspects such as changing market conditions, fluctuating production demand forecasts and 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 to high flexibility of the production-system focusses on 
realizing high availability of the technical systems for manufacturing purposes. The weight 
attributed by production to the availability of a technical system depends on its importance for 
the production process. Corrective maintenance of a technical system regarded to be essential has 
be carried out immediately and as fast as possible; Preventive maintenance of such a system has 
to be executed during periods dictated by production. The requirements of production in view of 
the availability of not-essential technical systems are not that rigid. The actual execution of 
corrective maintenance demand can be delayed for some time without problems; The execution 
of preventive maintenance demand can be tuned in with production on the short-term.

The production economic objectives imply that the maintenance capacities should be used 
efficiently. The limited capacities have to be shared by all the technical systems. This competition 
may result in waiting times for some of these systems. Furthermore, the efficient use of capacities 
may favour the execution of operations earlier than demanded to take advantage of common set-
ups. Both considerations may have an adverse effect on the eventual availability of the technical 
systems. At management control level, the conflict between availability on the one hand 
and efficiency on the other hand has to be resolved resulting in the operational constraints.

Operational constraints are structural conditions under which the maintenance unit should 
perform satisfactorily. These constraints primarily concern capacity aspects such as use, adjustment 
and utilization. To avoid unsolvable problems within the maintenance unit, the amount of capacity 
used for preventive maintenance has to leave room for corrective maintenance which will 
inevitably occur in due time. Furthermore, the mix of corrective maintenance load and urgency 
has to be under control to avoid disruption of planned preventive maintenance. Capacity 
adjustment is possible at certain prices and with certain leadtimes. Capacities can only be used 
efficiently if there is a certain amount of work simultaneously available. Workload norms i.e. 
upper and lower bounds have to be applied in view of the bottleneck capacities. Due to the 
inherent uncertainty of work content, the constraints take the form of rough guidelines.

Given maintenance concepts and maintenance inventory control, operational maintenance 
control essentially concentrates on the timing of the allocation of capacities to operations 
demanding them. In view of the operational constraints formulated at management control level, 
the control problem can be broken down into two sub-problems (fig. 3):

- work order coördination;
- maintenance unit control.

Work order coördination essentially concentrates on how to realize maintenance demand, the production requirements and the operational constraints simultaneously. Maintenance unit control is concerned with the problem how to achieve the agreed performance in terms of timing and efficiency of the work to be carried out, given that the operational constraints are met.

**WORK ORDER COORDINATION**

Work order coördination defines work orders on the basis of maintenance demand and production requirements, taking into account the operational constraints and the availability of material. A work order is a collection of operations to be carried out by the maintenance unit conform agreements in terms of timing and efficiency. Work order coördination consists of the following three decision functions (fig. 4):

- preventive maintenance planning;
- corrective maintenance classification;
- adaptive maintenance acceptance.

Preventive maintenance planning

Preventive maintenance planning deals with generating preventive maintenance work orders. This decision has to see to efficiently meeting preventive maintenance demand and the requirements of production, taking into account operational constraints and material availability.

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 view of preventive maintenance of essential technical systems dictate execution during the non-productive periods in the production pattern agreed upon at the aggregate production planning level. Maintenance of not-essential technical systems is not critical and 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. At this level only the period in which this work
has to be carried out is determined. Operational restrictions limit the size of work orders and the preventive maintenance load which may be planned. Capacity problems can be circumvented by making use of volume-flexibility. The overall capacity need of preventive maintenance may be smoothed anticipating 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 accepted by maintenance and have to be carried out conform planning. Consequently, this flow forms input for work order scheduling directly. The other flow consists of advancable work orders the ultimate acceptance of which still has to be decided upon.

**Corrective maintenance classification**

Corrective maintenance classification is concerned with the specification of corrective maintenance work orders. This decision has to reflect consequences of failure for production, actual state of the maintenance capacities and operational constraints.

In view of each failure, production and maintenance management have to reach an agreement about the urgency of execution. Production management will have to translate the importance of the system for their day-to-day operations into an acceptable duration of unavailability. Maintenance management has to take into account its possibilities which are the outcome of already accepted work and the operational constraints regarding the distribution of work in terms of urgency and load.

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. Consequently, these work orders are input of work dispatching directly. The other flow encompasses postponable work orders which are placed in a buffer awaiting acceptance.

**Adaptive maintenance acceptance**

Adaptive maintenance acceptance aims at smoothing the amount of work to be processed by the maintenance unit. Fluctuations are caused by the predictable variations in fixed maintenance and unpredictable variations in rush maintenance. Advancable and postponable work orders form a kind of buffer of potential work. To perform its function properly, the size of the buffer should remain between a minimum and a maximum value both dictated by operational constraints. Undershooting the minimum value activates preventive maintenance planning to advance demand additionally with a detrimental effect on efficiency of execution. Overshooting the maximum value signals the need for contracting-out maintenance. It has to be decided upon which work orders are going to be subcontracted.
Adaptive maintenance acceptance results in two distinct flows of work. One flow consists of work to be carried out externally. The other flow consists of work accepted to be carried out by the maintenance unit. These adaptable work orders are going into the portfolio of work which forms one of the three inputs for maintenance unit control.

MAINTENANCE UNIT CONTROL
Maintenance unit control focuses on realizing the work orders according to the agreements on timing and efficiency, taking into account the short-term production wishes. The following three decision functions can be distinguished (fig. 5):

- work order release;
- work order scheduling;
- work dispatching.

Work order release
Work order release has to control the amount of work in progress by means of setting free adaptable work orders from the portfolio of accepted orders to compete for the scarce capacities.

Each period the amount of work released should be proportional to the capacity available in the next period. During this period, however, the remaining workload may drop below a certain level e.g. because the amount of rush work is less than expected. Additional work can be released then. The decision which specific work orders from the portfolio are going to be released depends on such aspects as release opportunities, material availability and work order priorities. Release opportunities in terms of efficiency are unexpected production interruptions or possibilities to combine the execution of various orders thus reducing execution cost. 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 due date.

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 released adaptable 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
systems have to be accounted for as well as short-term capacity bottlenecks, environmental conditions and efficiency considerations. Capacity problems can be tackled by using volume-flexibility.

The eventual schedule specifies the start and finish dates or the period in which each distinguished work order has to be carried out if some slack still remains. It should be stressed that in view of the high degree of uncertainty inherent in maintenance, the aim of work order schedule is to set forth a sound basis for work dispatching. Generally it will not be possible to adhere completely to the schedule; problems in work progress may require re-scheduling of adaptable or even re-planning of fixed work orders.

**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 work progresses according to the work order schedule, sequencing of work orders to be executed by a specific capacity type has to take into account the relative priorities between the distinguished classes of demand. Within each class, sequencing can make use of appropriate priority rules. However, if progress lags behind then the relative priorities of the classes and the work order schedule are no longer a sound basis for decision making. The relative importance of the individual work orders determines the sequence in which they will be carried out. This importance is derived from aspects such as the essentiality of the technical system concerned and the remaining processing time.

In general, a work order will be allocated to a specific capacity type if that type becomes available. However, fixed and rush work orders may require interrupting work being carried out ("preemptive priority"). To avoid the waiting of work orders for a capacity type to become idle or ongoing work to be interrupted, it can be decided to make use of mix-flexibility, if available of course.

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 of the maintenance engineers.

**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 is available to counter the inevitable variations in demand. The forementioned characteristics favour 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 inevitable "variances". As new information arises continuously, even during execution of maintenance, feedback and progress monitoring are essential ingredients of any maintenance control problem.

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. 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 decisionmaker.

The major assumption in this paper is that demand originates either from the maintenance concept or from failures signalled by production. However, determination of the state of a technical system and altering this state can be considered to be defined as two distinct work orders. The outcome of the execution of diagnosis is then feedback to corrective maintenance classification, of inspection to preventive maintenance planning. Furthermore, a work order may be conditionally defined. Repair/recondition is directly carried out after diagnosis/inspection if execution takes less than a specified amount of time. If more time is needed then the information is feedback to work dispatching, work order scheduling and eventually to corrective maintenance classification/preventive maintenance planning. Which decision is going to deal with the demand at hand depends on slack remaining at each level. Another source of "bottom up" demand is related to "minimal repairs". A minimal repair aims at tying the technical system over until the repair can be done properly e.g. without unduly interrupting the production process or with the use of the right parts. Demand originating during execution of maintenance enhances the importance of a hierarchical control structure, of vertical information systems to feedback information and of the formulation of additional operational constraints to deal with the intricacies with respect to defining work orders. Minor assumptions are failures resulting in breakdown of the technical system, maintenance execution requiring the technical system to be unavailable, secondary damage playing no role and work orders being one-man jobs. 

11
assumptions do not affect the framework; They do complicate the decisions to be made.

Practical research is being directed at the applicability of (parts of) the framework in establishing operational maintenance control systems in real situations. This requires the design of organisable and implementable decision functions taking into account the characteristics of the situation to be controlled. Further theoretical research is needed into management control. It concerns topics such as integrating production and maintenance, coördination of maintenance concept design, maintenance inventory control and operational maintenance control and formulation of the operational constraints.

LITERATURE

Fig. 1: The relationship between production and maintenance
Fig. 2: Maintenance control
Fig. 3: The structure of maintenance control
Fig. 4: The structure of work order coordination.
Fig. 5: The structure of maintenance unit control