TECHNICAL
RESEARCH
REPORT

SYSTEMS
RESEARCH
CENTER

Supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), Industry and the University

Hierarchical Modeling Approach for Production Planning

by G. Harhalakis, R. Nagi and J.M. Proth

TR 92-14

Hierarchical Modeling Approach for Production Planning

G. Harhalakis[1], R. Nagi[1*], J.M. Proth[2]

[1]Department of Mechanical Engineering & Systems Research Center, University of Maryland, College Park, MD 20742
[2]INRIA, 4, rue Marconi, Technopole Metz 2000, 57070 METZ, France, & SRC, University of Maryland, U.S.A.

Production management problems are complex owing to large dimensionality, wide variety of decisions of varying scope, focus and time-horizon, and disturbances. A hierarchical approach to these problems is a way to address this complexity, wherein the global problem is decomposed into a series of top-down sub-problems. We advocate that a single planning architecture cannot be employed for all planning problems. We propose a multi-layer hierarchical decomposition which is dependent on the complexity of the problem, and identify the factors influencing complexity. A systematic stepwise design approach for the construction of the hierarchy and inputs required are presented. The subsequent operation of the hierarchy in an unreliable environment is also explained. Aggregation schemes for model reduction have been developed and blended with a time-scale decomposition of activities to provide the theoretical foundation of the architecture. It is also hoped that this methodology can be applied to other such large-scale complex decision making problems.

1 Introduction

Most real-world production management problems are complex owing to large dimensionality, wide variety of decisions of varying scope, focus and time-horizon, and disturbances. Time-horizon refers to the length of time over which decisions are performed. Disturbances could be both endogenous (e.g., resource failures) as well as exogenous (e.g., unscheduled orders) random events. This complexity warrants hierarchical approaches, wherein the global problem is decomposed into a series of sub-problems. These sub-problems are sequentially solved in a top-down manner; the global solution is obtained when all the problems are solved. The principles common to these approaches are that the higher levels are more aggregate, with longer horizons, whereas, the lower levels are more detailed, with shorter horizons.

Apart from the reduction of complexity, another important benefit of hierarchical decision making is obtained when the system is subject to random events. Monolithic models would require the entire problem to be resolved, while the hierarchical approach can gradually absorb random events without the need to resolve higher level problems. This results in large savings in computational burden. Decisions at various levels in the planning process are made at different points in time. Higher level decisions are more aggregate, and need not explicitly consider uncertain data at detailed levels. Random events with a lesser impact on the system can be absorbed at lower levels. Hierarchies allow reduction of detailed information and longer "look ahead" capabilities. In addition, forecasting is usually easier and more accurate for aggregates than for detailed entities.

Literature surveys in the field of hierarchical production management can be found in [2,3,7]. These works consider different sets of assumptions, layering methods, number of layers and decision making methodologies. The work of Hax and Meal [5] has been considered a substantial contribution. While the Hax-Meal framework has several advantages, some of the shortcomings are: (i) the product aggregation

---

* Author to whom correspondence should be addressed.

scheme is only relevant to a particular class of systems (models are based on a typical cost structure), (ii) no randomness is taken into account, and (iii) no spatial aggregation of the system is proposed.

Production planning problems differ in nature and complexity. In this view, it seems that no single planning architecture can be employed for all planning problems. The architecture of the planning system, the number of levels in the hierarchy, the criteria of relevance, and the control to be applied are problem dependent. This paper aims at defining a modeling approach for such hierarchical architectures, which are dependent on the complexity of the problem. A framework for designing a hierarchical system is proposed. We also present in brief, the execution of such a system to effectuate optimal decision making.

1.1 Factors influencing problem complexity

The factors influencing the nature and complexity of production planning problems depend on the manufacturing system under consideration as well as the demand characteristics. Issues relating to the manufacturing system concern : i) dimensionality (number of machines and parts), ii) type/characteristics of products and production methods (mass, batch, or jobbing production), and iii) disruptive stochastic events (work-center failures). Issues associated with the demand characteristics are: i) dynamics of the demand, e.g. trends, seasonality, ii) stochastic nature of demand, and iii) accuracy, nature and horizon of forecast. Changes in product mix, frequent order cancellations, changes, seasonality, and trend of the demand are the demand dynamics that effectuate complexity in the planning process. Thus, it is apparent that hierarchical planning architectures are highly influenced by the specific problem characteristics.

1.2 Issues relating to hierarchical architectures

While such hierarchical methods have been applied to production systems in the past, several related issues remain unanswered. The essential issues that draw attention in a hierarchical architecture are: (i) the construction of the hierarchy (number of levels, models and horizons) depending on the problem and its complexity, (ii) controllability, i.e., the optimal control at a given level leads to a set of constraints defining a non-empty set of feasible controls at the next lower level; thus the system can be directed to a desired state by the top-down constraint propagation procedure, and (iii) consistency. In the case of similar criteria at two levels, consistency means that the optimum of the global problem lies within the set of feasible controls at the lower level generated as a consequence of the high level set of constraints; thus the optimal solution of the individual problems results in the optimal solution of the global problem.

The paper is organized as follows. In section 2, the methodology for design, detailing the structural, aggregation, and decomposition issues, is presented. Section 3 details the inputs required in the design process. The design process is described in section 4. Section 5 is devoted to the operation of the hierarchy. Finally, in section 6 we draw our conclusions and present our future work.

2 Methodology of Hierarchical Design

The objective of this section is to present a methodology for the construction of such a hierarchical system. The present effort is limited to the construction, and the subsequent use of the architecture for decision making. Some consistency issues relating to a set of criteria have been addressed in Nagi [9]. However, for general and multi-criteria problems, the consistency issue is difficult to demonstrate.

2.1 Structural Issues

Consider a n-level hierarchy. Each level in the hierarchy is described by a model and an associated Decision Making Problem (DMP). A model consists of a set of entities. Each entity is associated with a set of attributes. Each attribute can be assigned a value from a value set associated with it. Decision making consists of determining a set of optimal controls from a set of feasible controls, such that, the constraints specified are satisfied over a horizon, and some criteria are optimized. Thus, it is required to specify: (i) a set of constraints (from upper level), (ii) a set of feasible controls, (iii) a set of criteria, and (iv) a horizon.

Note that the DMP and the model at each level are highly related. For example, the values of some attributes of some entities should be modifiable within the horizon by the application of a feasible control.

2.2 Aggregation Issues

Aggregation-Disaggregation is another major issue concerned with the design of hierarchical systems. Krajewski and Ritzman [6] provide a survey of the problems and research in this field. Aggregate production planning is essentially performed to decide resource/work force levels. Translating demand forecasts for a wide range of products into resource requirements is a difficult task, which is further complicated by the uncertainty of demand forecasts. The aggregate production planning methodology can be applied to any level of the hierarchy; the higher levels in the hierarchy are more aggregate. Very often, spatial aggregation, i.e. aggregation of production facilities is also performed. What the appropriate aggregating schemes should be is not always obvious. The scheme is very often chosen based on a typical cost structure, [5]. From the point of view of resource level requirements, it is natural to consider products having similar processing requirements, Meier [8]. However, in light of stiffer industrial competition, better planning and resource allocation is becoming increasingly important.

Nagi [9] presents the underlying aggregation theory. A two-level hierarchy is developed for holding and backlogging costs employing this aggregation scheme, and optimality is demonstrated in a particular case. In the general case, we summarize the product and machine aggregation as follows. Products are aggregated into families by the K-mean algorithm in cluster analysis. Product entities are represented in IR^{m+c} by a point. m axes represent the processing time required by the product entity on the m machine entities, and c axes represent the attributes relevant to the criteria (at that level). For instance if the criteria are earliness and tardiness, c equals 2, and the axes represent per unit holding cost and backlogging cost, respectively. Each point is weighted by the long term production volume of the corresponding product. Then, the K-mean algorithm is employed to determine the clusters or product families and their attributes. The advantages of aggregating parts into families in this manner allow for some uncertainty in the demand to be absorbed while reducing variances. Furthermore, it reduces the level of detail required for future production periods, i.e. forecasts in terms of aggregates are sufficient, and detailed product forecasts are not required. Aggregation of machines into cells is also performed, [9]. In addition, temporal aggregation is also performed.

2.3 Decomposition principles

In this section, we present the principles that are employed in decomposing the overall problem into sub-problems, i.e. in the construction of the hierarchy. Time-scale decomposition is a technique developed for the analysis of dynamic systems in which different components of the state vector have very different dynamics. In this decomposition, the modes of the system are partitioned into classes, in such a way that each class is either fast or slow, with respect to the other classes, [1,10]. The literature in control theory essentially treats *multi-level* hierarchies. Unfortunately, this technique has not been developed substantially in the *multi-layer* literature. Gershwin [4], employs the frequency separation principle for hierarchical decomposition of FMS scheduling. The essential idea is to treat quantities that vary slower as static and to model quantities that vary faster in a way that ignores the details of their variations (considering averages).

In our methodology, we employ similar concepts of time-scale or frequency decomposition, blending it with the aggregation aspects at the higher levels of the hierarchy. This is intended to also address the planning related hierarchy, that requires a broader spectrum of activities (of different time-scales) to be considered. Higher levels of the hierarchy are more aggregate, allowing for longer horizons and elementary periods associated with them. Thus, in our multi-layer structure for hierarchical production management systems, the controller is decomposed into algorithms (levels) operating at different time intervals. Higher levels control the slower aspects of the system (i.e. address activities of longer duration), they intervene less frequently, with longer optimization horizons, and are based on more aggregate models. Progressively, the lower levels address faster aspects of the system over shorter optimization horizons (and associated elementary periods), while becoming more detailed.

While the design of the hierarchical controller is based on controllable activities of different time-scales, it is also intended to address uncontrollable activities. These uncontrollable activities (random events) are of different time-scales too. Each level of the controller treats activities (controllable/uncontrollable) with longer durations as static, and treats activities that vary faster by representing them as averages. Finally, the activities of comparable durations are addressed at a particular level. The controllable activities are planned at this level (controls), while the uncontrollable ones are absorbed. This decomposition of activities or controls also directly impacts on the calculation of the attributes of the entities at the different levels of the hierarchy.

3 Inputs to the Design Process

The design of the planning hierarchy requires a variety of information specifying the manufacturing system as well as the managerial goals and decisions. In this section, we present a list of characteristics (see [9] for details) that attempts to encompass most manufacturing and management systems.

3.1 Manufacturing system details

1) Work-center details : These consist of the number and types of work-centers in the system. If they are prone to failure, then the various failure modes and associated time distributions are of relevance. Given several failure modes of each machine, based on its Mean Time To Repair (MTTR), machines can be classified along a time-scale, where each group is represented by a characteristic MTTR.

2) Worker details : Worker hiring, firing, regular and overtime costs per unit of time are relevant in the case of labor intensive production. If the workers default, then the various default modes and associated distributions of time between absenteeism and time of absence are of relevance.

3) Product details : These consist of the set of product types, and their Bills-Of-Materials. For a unit of each product, the cost for holding it per unit time, and the penalty cost for backlogging, are important.

4) Routing details : These consist of the sequence of operations, set-up and run times.

5) Historical or Forecast data : Long term production requirements (or demand) of products.

3.2 Managerial inputs

1) Controls, U^0 : Controls are decisions that can be performed during the production planning process. Decision to sub-contract products, hiring or firing, decisions of overtime levels are related control. Performing set-ups, or loading parts are controls performed more frequently than those mentioned earlier.

Each control is associated with a period or duration of time which is elapsed before the results of the control can be observed (or expected). This duration depends on the type of control, and the speed of the system to accomplish the action. This duration is referred to as the response time of the system to this control. For instance, production of a product has a response time of the order of its cumulative lead time. Hiring of new workers is a control which generally has a longer response. Thus, each control can be assigned an order of magnitude for its response time in the particular manufacturing system.

2) Criteria, C^0 : Criteria are the management objectives that have to be optimized during production planning. Minimization of production costs, distribution costs, inventory costs, backlogging costs, hiring and firing costs, regular and overtime labor costs, and set-up costs may be relevant. It is important to indicate that the criteria bear a strong relationship with the controls considered.

4 Design of the Planning Hierarchy

In this section, we present the design procedure for the construction of a planning hierarchy, performed in a bottom-up manner. This design procedure requires significant human interaction and decision making.

We begin at the bottom level (level 1), and consider the physical entities, e.g. products and machines. Therefore, the model is the physical one (and known), and we need to define the DMP. The elementary period, Δ^1 is determined based on several factors that include product lead times, intervals between shipment and inventory updates, to mention a few. For instance, if products have lead-times of the order of 10-15 minutes, and the inventory is updated daily, then one day is an appropriate elementary period. Since the elementary period at this level is usually of a fairly short duration, not all controls can be effectuated during it. Thus, only a set of controls that have a response time at least an order of magnitude less than Δ^1 are considered. This elementary period is also intended to absorb random events having a response time of the order of these controls.The capacities of the resources are represented in a manner that subtracts an amount equal to the expected durations of non-productive activities (random events like failures) occupying the resources, which have mean response times much less than Δ^1.

This set of controls defines the complexity of the DMP (number of variables and constraints), after which the length of the horizon can be ascertained either by the complexity of the problem or the solution time permissible. Thus, owing to detailed nature and large dimension of level 1, the horizon is limited by the complexity of the problem. Finally, the criteria to be considered are determined based on the controls at this level. This completes the definition of the DMP at the bottom level. Note at this time the next higher level is not known, hence the high level constraints to this problem are not know exactly, although their general form can be speculated.

The need to address planning over a longer horizon than the current one, as well as the need to consider all possible controls requires to proceed to develop higher levels. The essential idea is to reduce the dimension and complexity of the current model by temporal and entity aggregation. We employ the theory developed in Nagi [9] to accomplish these aggregations in a manner that is consistent with the criteria at the bottom level. The resources are also aggregated, and their capacities are represented to reflect the absorption of the random events at level 1 as well. Following this aggregation, the attributes of the newly derived entities are computed. In this manner, we develop the model at the next upper level (level 2).

Having determined the model and entities of the level under consideration, we need to define the DMP. Once again, we select the appropriate elementary period, Δ^2 which is based on the following guide-lines : it is usually a multiple of Δ^1, and is no greater than the horizon at the lower level. Since the elementary period at this level is of longer duration than that of the lower level, some additional controls can be effectuated during it. The earlier (lower level) controls are also considered, related now to the aggregate entities. This set of controls then defines the complexity of the DMP, following which the horizon is ascertained.

The subsequent higher levels are designed in a similar fashion, i.e. aggregating and defining the model followed by defining the DMP. The process is repeated until the desired horizon is obtained and all controls have been addressed.

5 Operation of the Planning Hierarchy

Once the hierarchy is constructed, it can be employed to solve the planning problem. Given that the DMPs at each level and the associated solution algorithms are not defined here precisely our intent is to conceptually present the overall functioning of the hierarchy.

5.1 Top-down solution procedure to the planning problem

The solution procedure of the planning problem begins at the top-most level of the hierarchy, say level n. This level is the highest, hence the upper level constraints, T^n, are strategic constraints in this case. The demand at this level is expressed in terms of the aggregate product entity at level n. For the initial period(s), the demand can be taken from the bottom up aggregation of detailed customer orders, while for the latter periods it is forecasted. This level solves its DMP over the horizon H^n in order to optimize the criteria C^n. The solution of this problem X^n, truncated over the lower level horizon H^{n-1}, is then transmitted to the lower level (n-1). This is in the form of its upper level constraints T^{n-1} for level n-1. The primary objectives of level n-1 are: (i) the disaggregation of this aggregate production, and (ii) the computation of controls at this level, i.e. controls belonging to U^{n-1}, such that the criteria C^{n-1} are optimized over H^{n-1}. The production planning of product entities at each level is performed under capacity constraints of the corresponding machine entities. The top-down solution procedure continues for lower levels of the hierarchy in a similar manner. The final result of this top-down computation is the product production plan in the elementary periods Δ^1 over the horizon H^1, under resource capacity constraints. The hierarchy also employs a rolling horizon mechanism in order to progressively take future information into account, [9].

5.2 Reaction of the hierarchy to random events

In practice, there are a number of disturbances and random events that make it difficult to respect higher level decisions. This calls for a bottom-up feedback procedure in that each level transmits the difference between the planned and accomplished states to the next higher level. Usually, such a feedback procedure originates at the bottom level of the hierarchy due to random events on the shop floor, and is transmitted upwards until it can be absorbed by some level. However, there are certain other random events in the production planning environment that have a response time comparable to the duration of elementary periods at some level of the hierarchy. For example, a worker strike, or a severe machine breakdown can have durations comparable to durations of a high level elementary period of the hierarchy. In this case, when such a random event arrives in the system, the response time of the system it is evaluated, and addressed at an appropriate (higher) level of the hierarchy.

6 Conclusions

We propose a multi-layer hierarchical decomposition approach to complex large scale production management problems. The approach is generic, and is not based on a fixed aggregation scheme or number of levels; it depends on the characteristics of the production system and the complexity of the DMP at hand. A systematic stepwise design approach for the construction of the hierarchy is presented. The subsequent operation of the hierarchy is also explained. The aggregation/disaggregation techniques for model reduction have been blended with a time-scale decomposition of activities. The aggregation schemes are hoped to provide consistent solutions, as the worst case analysis (Nagi [9]) of the hierarchical approach supports its applicability. The approach has been applied to a sample problem, and some numerical simulations are underway.

References

[1] Chow, J.H. and Kokotovic, P.V.: "Time Scale Modeling of Sparse Dynamic Networks," IEEE Transactions on Automatic Control, Vol. AC-30, No. 8, 1985.

[2] Dempster, M.A.H., Fisher, M.L., Lageweg, B., Jansen, L., Lenstra, J.K. and A.H.G. Rinnoy Kan: "Analytical evaluation of Hierarchical Planning Systems," Operations Research, Vol. 29, No. 4, 1981.

[3] Gelders, L.F. and van Wassenhove, L.N.: "Hierarchical Integration in Production Planning : Theory and Practice," Journal of Operations Management, Vol. 3, No. 1, 1982.

[4] Gershwin, S.B.: "Hierarchical Flow Control : A Framework for Scheduling and Planning Discrete Event Manufacturing Systems," IEEE Proceedings : Special Issue on Discrete Event Systems.

[5] Hax, A.C. and Meal, H.C.: "Hierarchical Integration of Production Planning and Scheduling," in Studies in the Management Sciences, M.A. Geisler, ed., Vol. 1, Logistics, North Holland - American Elsevier, 1975.

[6] Krajewski, L.J. and Ritzman, L.P.: "Disaggregation in manufacturing and service organizations: survey of problems and research," Decision Sciences, Vol. 8, No. 1, 1977.

[7] Libosvar, C.: "Hierarchical Production Management: The Flow Control Layer," Ph.D. Thesis, Univ of Metz, France, '88.

[8] Meier, K.: "Commande Hierarchisee d'un Systeme de Production," Ph.D. Thesis, University of Metz, France, '89.

[9] Nagi, R.: "Design and Operation of Hierarchical Production Management Systems," Ph.D. Thesis, Univ of MD, '91.

[10] Sandell, N.R., Varayia, P., Athans, M.A. and Safonov, M.: "A Survey of Decentralized Control Methods for large Scale Systems," IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, 1978.