

DISSERTATION DEFENSE

Emre Nadar

Thursday, April 26th, 2012

10:45 am

324 GSIA (West Wing)

New Markov Decision Process Formulations and Optimal Policy Structure for Assemble-to-Order and New Product Development Problems

This thesis examines two complex, dynamic problems by employing the theory of Markov Decision Processes (MDPs). The first two chapters consider assemble-to-order (ATO) inventory systems. An ATO system consists of several components and several products, and assembles products as demand is realized; it is becoming increasingly popular since it provides greater flexibility in manufacturing at a reasonable cost. This work contributes to the ATO research stream by characterizing optimal inventory replenishment and allocation policies. The last chapter examines new product development (NPD) with scarce resources and many projects in parallel, each lasting several periods in the face of uncertainty. This study advances the NPD literature by revealing that optimal project selection and resource allocation decisions are congestion-dependent. Below, we elaborate on the novel optimal policies and structural results we obtain using MDP formulations, which is the overarching theme of each chapter.

In the first chapter, we consider generalized ATO “ M -systems” with multiple components and multiple products. These systems involve a single “master” product which uses multiple units from each component, and multiple individual products each of which consumes multiple units from a different component. An example of such a system would be a manufacturer who sells an assembled product as well as individual spare parts. We model these systems as an infinite-horizon MDP under the discounted cost criterion. Each component is produced in batches of fixed size in a make-to-stock fashion; batch sizes are determined by individual product sizes. Production times are independent and exponentially distributed. Demand for each product arrives as an independent Poisson process. If not satisfied immediately upon arrival, these demands are lost. Therefore the state of the system can be described by component inventory levels. A control policy specifies when a batch of components should be produced (i.e., inventory replenishment), and whether an arriving demand for each product should be satisfied (i.e., inventory allocation). Since the convexity property that has been largely used to characterize optimal policies in the MDP literature may fail to hold, we introduce new functional characterizations for submodularity and

supermodularity restricted to certain lattices of the state space. The optimal cost function satisfies these new characterizations: The state space of the problem can be partitioned into disjoint lattices such that, on each lattice, (a) it is optimal to produce a batch of a particular component if and only if the state vector is less than a certain threshold associated with that component, and (b) it is optimal to fulfill a demand of a particular product if and only if the state vector is greater than or equal to a certain threshold associated with that product. We refer to this policy as a lattice-dependent base-stock and lattice-dependent rationing (LBLR) policy. We also show that if the optimization criterion is modified to the average cost rate, LBLR remains optimal.

The first chapter makes three important contributions. First, this study is the first to characterize the optimal inventory replenishment and allocation policies for M -systems. Second, this study is the first to characterize the optimal policies for any ATO problem when different products may use the same component in different quantities. Third, we introduce new functional characterizations restricted to certain lattices of the state space, giving rise to a lattice-dependent policy.

In the second chapter, we evaluate the use of an LBLR policy for general ATO systems as a heuristic. We numerically compare the globally optimal policy to LBLR and two other heuristics from the literature: a state-dependent base-stock and state-dependent rationing (SBSR) policy, and a fixed base-stock and fixed rationing (FBFR) policy. Taking the average cost rate as our performance criterion, we develop a linear program to find the globally optimal cost, and Mixed Integer Programming formulations to find the optimal cost within each heuristic class. We generate more than 1800 instances for the general ATO problem, violating the M -system product structure. Interestingly, LBLR yields the globally optimal cost in all instances, while SBSR and FBFR provide solutions within 2.7% and 4.8% of the globally optimal cost, respectively. These numerical results also provide several insights into the performance of LBLR relative to other heuristics: LBLR and SBSR perform significantly better than FBFR when replenishment batch sizes imperfectly match the component requirements of the most valuable or most highly demanded product. In addition, LBLR substantially outperforms SBSR if it is crucial to hold a significant amount of inventory that must be rationed.

Based on the numerical findings in the second chapter, future research could investigate the optimality of LBLR for ATO systems with general product structures. However, as we construct counter examples showing that submodularity and supermodularity, which are used to prove the optimality of LBLR in the first chapter, need not hold for general ATO systems, showing optimality of LBLR for general ATO systems will likely require alternate proof techniques.

In the third chapter, we study the problem of project selection and resource allocation in a multi-stage new product development (NPD) process with stage-dependent resource constraints. As in the first two chapters, we model the problem as an infinite-horizon MDP under the discounted cost criterion. Each NPD project undergoes a different experiment in each stage of the NPD process; these experiments generate

signals about the true nature of the project. Experimentation times are independent and exponentially distributed. Beliefs about the ultimate outcome of each project are updated after each experiment according to a Bayesian rule. Projects thus become differentiated through their signals, and all available signals for a project determine its category. The state of the system is described by the numbers of projects in each category. A control policy specifies when and at what rate to utilize the resources at each stage, and on which projects.

We characterize the optimal control policy as following a new type of strategy, state-dependent non-congestive promotion (SDNCP), for two different special cases of the general problem: (a) when there is a single informative experiment, or (b) when there are multiple uninformative experiments. An SDNCP policy implies that, at each stage, it is optimal to advance a project with the highest expected reward to the next stage if and only if the number of projects in each successor category is less than a state-dependent threshold. Specifically, threshold values decrease in a non-strict sense as a later stage becomes more congested or as an earlier stage becomes less congested; a stage becomes more congested with an increase in the number of projects at this stage or with an increase in the expected reward of any project at this stage. We also report the strong numerical performance of an SDNCP policy as a heuristic for the general problem. Taken together, these findings highlight the importance of congestion in optimal portfolio strategies.