

DISSERTATION PROPOSAL

Emre Nadar

Monday, February 6th, 2012

9:30 am

324 GSIA (West Wing)

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

My dissertation examines two complex, dynamic problems by employing the theory of Markov Decision Processes (MDP). In the first two chapters, I examine 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. I contribute to the ATO research stream by characterizing optimal inventory replenishment and allocation policies. In the last chapter of my dissertation I study the MDP representation of a challenging Operations Management problem that includes dynamics and uncertainty. This chapter considers new product development (NPD) with scarce resources and many projects in parallel, each lasting several periods in the face of uncertainty. In two special cases, I show that the optimal policy for project selection and resource allocation is congestion dependent. Below, I elaborate on the novel optimal policies and structural results I obtain using MDP formulations, which is the overarching theme of each chapter.

Chapter 1: New Functional Characterizations and Optimal Structural Results for Assemble-to-Order *M*-Systems

In the first chapter, I 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.) I 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, I 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. I refer to this policy as a *lattice-dependent base-stock* and *lattice-dependent rationing* (LBLR) policy. I 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 attempt 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, I introduce new functional characterizations restricted to certain *lattices* of the state space, giving rise to a *lattice-dependent* policy.

Chapter 2: Performance Evaluation of Lattice-Dependent Base-Stock and Rationing Policies for Assemble-to-Order Systems

In the second chapter, I evaluate the use of an LBLR policy for general ATO systems as a heuristic. I 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, I develop a linear program to find the globally optimal cost, and Mixed Integer Programming formulations to find the optimal cost within each heuristic class. I 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 I 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.

Chapter 3: Optimal Portfolio Strategies for New Product Development

In the third chapter, I 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, I 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.

I characterize the optimal control policy as following a *new* type of strategy, *state-dependent non-congestive promotion*, for two different special cases of the general problem: (a) when there are multiple uninformative experiments, or (b) when there is a single informative experiment with multiple signals. A *non-congestive promotion* 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. This result uncovers the role congestion plays in optimal policies: 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. The third chapter concludes with a conjecture that the optimal control policy for the general problem is a state-dependent non-congestive promotion policy.

My dissertation widens our knowledge of optimal policies for MDPs by advancing novel structural results for the challenging ATO and NPD problems under Markovian assumptions. To our knowledge, the first chapter is the first attempt to characterize optimal policies for the M -system. Computational results in the second chapter further emphasize the practicality of an LBLR policy for the general ATO problem. In the last chapter, I establish the optimality of the state-dependent non-congestive promotion policy in two special cases of the NPD problem. Verification of the optimality of state-dependent non-congestive promotion for the general NPD problem would have substantial implications for many industries including, but not limited to, pharmaceutical and IT where R&D plays a vital role. For this purpose, during the rest of my Ph.D. studies, I intend to conduct numerical experiments to evaluate the use of this policy as a heuristic. I also plan to numerically compare it to other simpler policies, such as a fixed non-congestive promotion policy. Such a computational study may reveal surprising results similar to those obtained with an LBLR policy on the general ATO problem.