3rd Rutgers Applied Probability Conference

Stochastic Models and Algorithms for Intelligent Business Systems

Saturday, June 7, 2014

100 Rockafeller Road - Room 1144, Piscataway, NJ 08854

In Honor of Eugene A. Feinberg

Breakfast. 8:00-8:30 AM

Opening Remarks. 8:30 - 8:40 AM, Dimitris N. Metaxas, Rutgers University and Joseph Mitchell, Stony Brook University

Session 1. 8:40 - 10:15 AM. Melamed Benjamin - Rutgers University, Chair

Berge's Maximum Theorem for Non-compact Image Sets and Its Applications.

Pavlo Kasyanov, Kyiv Polytechnic Institute & World Data Center, Ukraine and Eugene A. Feinberg, Stony Brook University.

For an upper semi-continuous set-valued mapping from one topological space to another and for a lower semi-continuous function defined on the product of these spaces, Berge's theorem states lower semi-continuity of the minimum of this function taken over the image sets. It assumes that the image sets are compact.

For metrizable topological spaces, we provide the extension of Berge's theorem to set-valued mappings with possible noncompact image sets, provide relevant properties of minima, and consider its applications for Markov Decision Models (MDPs) with incomplete information.

Cesaro Limits of Actions and Ergodic Control of Markov Chains.

Adam Shwartz, Technion, Israel & Jacobs Technion-Cornell Innovation Institute, Amarjit Budhiraja, University of North Carolina at Chapel Hill and Xin Liu, Clemson University.

We consider the ergodic control of discrete time controlled Markov chains with a locally compact state space and a compact action space. The first question we consider is: what characterizes general policies which have the same cost (for a large class of immediate cost functions) as a given stationary Markov policy?

A flexible family of controls, called *action time sharing* (ATS) policies, associated with a given continuous stationary Markov control, is introduced. These policies where first introduced by Altman and Shwartz for the case of countable state space and discrete control space. It is shown that the long term average cost for such a control policy, for a broad range of one stage cost functions, is the same as that for the associated stationary Markov policy.

ATS policies are well suited for a range of estimation, information collection and adaptive control goals. To illustrate the possibilities we present two examples: The first demonstrates a construction of an ATS policy that leads to consistent estimators for unknown model parameters while producing the desired long term average cost value. The second example considers a setting where the target stationary Markov control q is not known but there are sampling schemes available that allow for consistent estimation of q. We construct an ATS policy which uses dynamic estimators for q for control decisions and show that the associated cost coincides with that for the unknown Markov contol q.

Average Cost Markov Decision Processes.

Eugene A. Feinberg, Stony Brook University.

This talk describes some of the major developments in the theory of Markov Decision Processes with average costs per unit time over the last 55 years: from models with finite state and action sets to the models with uncountable state and action sets that cover applications to periodic review inventory control systems.

Break.

Session 2. 10:30 - 12:05 AM. David.Hunt Oliver Wyman, Chair

Risk neutral and risk averse approaches to multistage stochastic programming.

Alexander Shapiro, Georgia Institute of Technology.

In many practical situations one has to make decisions sequentially based on data available at the time of the decision and facing uncertainty of the future. This leads to optimization problems which can be formulated in a framework of multistage stochastic programming. In this talk we consider risk neutral and risk averse approaches to multistage stochastic programming. We discuss conceptual and computational issues involved in formulation and solving such problems. As an example we give numerical results based on the Stochastic Dual Dynamic Programming method applied to planning of the Brazilian interconnected power system.

Structures of Optimal Policies in MDPs with Unbounded Jumps: the State of our Art.

Floske Spieksma, University of Leiden, Netherlands and Herman Blok, University of Leiden, Netherlands.

A standard method for the derivation of structural results for the optimal policy is to show that the desired structural properties propagate through a value iteration step. This requires the Markov Decision Process to be uniformisable, hence to have bounded jumps. However, many modern applications of Markov processes require unbounded jump intensities as a function of state.

Stony Brook University: Department of Applied Mathematics and Statistics
informs - Institute for Operations Research and the Management Sciences, New Jersey Chapter
Rutgers: Departments of Management Science & Information Systems and of Supply Chain Management & Marketing Sciences
The Center for Dynamic Data Analytics, the Rutgers Center for Supply Chain Management, and the Rutgers Center for Operations Research

Under mild drift and continuity conditions, structural results for uniformisable perturbations carry over to the original non-uniformisable model in the case of the expected discounted cost criterion. For the expected average cost criterion, similar results hold, although under a quite restrictive drift function criterion implying exponential ergodicity (in norm). If the property that each stationary policy induces an ergodic Markov process does not hold, there is no satisfactory theory so far that validates deriving structural properties of optimal policy by an analysis of suitable uniformisable perturbations.

We discuss the various drift conditions used in the above mentioned analyses, as well as an application to a web farm model. In this model, there are policies for which the associated Markov process is transient. However, it has some interesting features that do allow to validate a limit argument for a sequence of uniformisable perturbations.

Preferences, Risk-Neutrality, and Risk-Sensitive MDPs.

Matthew J. Sobel, James Alexander, Case Western Reserve University.

Four axioms imply that a binary preference relation among stochastic processes corresponds to a binary preference relation among present values (which are random variables). Risk sensitivity in discounted Markov decision processes (MDPs) is often modeled with the expected value of an exponential function of a risk parameter times the present value of a time stream of random monetary amounts. What condition ensures that a non-zero exponential risk parameter is logically consistent with discount factors in present values? This talk precisely answers the question embedded in the general setting of a binary preference relation on a real vector space. The condition is the four axioms without the converse of one of the four (the decomposition axiom).

Lunch Break.

Session 3. 1:30 - 4:00 PM. Michailidis George University of Michigan, Chair

Risk Management, Dynamic Resource Allocation, and Time-Cost Tradeoff in Projects Subject to Uncertainty.

Lynn Kuo, Manuel Nunez, University of Connecticut and Robert Chiang, Fordham University School of Business.

In the traditional project crashing problem (also known as the time-cost tradeoff problem) associated with a project network, the decision variables are assumed to be continuous, representing the amount of reduction in the duration of the tasks in the project, and the objective is to find the least expensive set of time reductions (crashing times) to complete the project within a given time limit (to avoid penalties). We will study a variant of this problem with a combinatorial probabilistic version, where the objective is to find the set of crashing choices that minimizes the time reduction cost and the expected penalty risk from tardiness. We show this problem is NP-hard, so we propose efficient heuristics that can provide approximate solutions to this problem. We also extend this problem to situations where revising resource assignment decisions are needed after reviewing the status of the project at several time points before the project deadline.

Change Point Inference for Time-varying Random Graphs.

George Michailidis, University of Michigan.

We investigate a model of a random graph, where the edges can be in a present/absent state. The states of each edge evolve as a Markov chain independently of the other edges, and whose parameters exhibit a change-point behavior in time. We derive the maximum likelihood estimator for the change-point and characterize its distribution.

Depending on a measure of the signal-to-noise ratio present in the data, different limiting regimes emerge. Nevertheless, a unifying adaptive scheme can be used in practice that covers all cases. We illustrate the model and its flexibility on US Congress voting patterns using roll call data.

Adaptive Monte-Carlo via Bandit Allocation.

Csaba Szepesvári, James Neufeld, Andras Gyorgy, and Dale Schuurmans, University of Alberta, Canada.

We consider the problem of sequentially allocating samples between a set of unbiased Monte-Carlo estimators with unknown variance. To minimize the mean-squared-error (MSE) of the final estimate, we develop strategies that are guaranteed to approximate the MSE of the single best estimator chosen in retrospect. We achieve these results by reducing sample-allocation to a stochastic multi-armed bandit problem, which allows well developed allocation strategies and corresponding analyses to be applied. These developments are then extended to a scenario where alternative estimators have different costs, yielding novel allocation strategies with strong guarantees. Overall, our new adaptive Monte-Carlo methods provide stronger guarantees than alternatives, while offering empirical improvements.

Adaptive sampling policies under incomplete information and non stationary distributions.

Apostolos N. Burnetas, University of Athens, Greece and Michael N. Katehakis, Rutgers University.

We consider the problem of adaptive sampling from a finite number of independent statistical populations under incomplete information regarding the underlying distributions. One population is selected for sampling at every period. The outcome of each period is scaled by a time-dependent factor and the objective is to maximize the expected sum of scaled observations. We adopt a frequentist framework and formulate the problem of minimizing the asymptotic rate of increase of an appropriately defined regret. We derive

a sufficient condition for the existence of consistent sampling policies and present some partial results on efficient policies and the optimal convergence rate. We also discuss relationships with the discounted multi-armed-bandit problem.

Decomposition Algorithms for Constrained MDPs with an Application to Wastewater Treatment Plants' Optimization.

Alexander Zadorojniy, **IBM Research - Haifa, Israel**, Adam Shwartz, **Technion, Israel & Jacobs Technion-Cornell Innovation Institute**, Segev Wasserkrug, Sergey Zeltyn, **IBM Research - Haifa, Israel.**

Our work is motivated by the need to optimize the operations of wastewater treatment plants (WWTPs). Many such plants consist of three main parts connected in series, with possible feedbacks between the parts. The sewage (or influent) enters the liquid line part of the plant, which outputs both treated water which must meet regulatory permits, and a more solid mass called sludge. The sludge then goes to the sludge line part of the plant, which treats the solids (sludge) to a level where they can either be used (e.g., as fertilizer), or safely disposed of. The liquids output by the sludge line are recycled back to the input of the liquid line. Another output produced by the sludge line is methane gas, which is passed on to the third part of the plant - the gas line. The gas line either stores the gas, or transforms it to electricity.

One main challenge is to operate the WWTPs in the most cost effective manner, while maintaining regulatory permits regarding the treated wastewater and sludge. In each such plant, there are numerous control options. In addition, the input to the plant (the influent) is highly dynamic with high daily variations, and the processes in the plant are complex biological, physical and chemical processes, usually modelled as a set of first order differential equations. Finally, while the effect of some control actions may be felt within minutes or hours, the effects of others, especially in the sludge line, may only be felt in a manner of days.

Constrained Markov Decision processes (CMDPs) are an ideal mechanism to model the operational optimization of such plants, as they can both capture the uncertainties inherent in the plant input, and abstract away the operational complexities by using a probability space representation. However, the state and action spaces of the whole plant are enormous (thousands of control actions and billions of states). To deal with this size we developed two algorithms which use a decomposition technique to reduce both the overall state space and the overall action space to several separate state and action spaces, each composed of several thousand and several hundred possibilities, of states and actions respectively. In both algorithms each part of the plant is modeled as a separate CMDP with a different state and action space. Both algorithms also propagate the costs back from the gas line, through the sludge line, and, finally, to the liquid line, in order to take into account, for example, the effects of actions taken by the liquid line on the costs of the sludge line. In addition, both algorithms use specialized constraints in order to take into account the effect the actions may have on the possible feedbacks between the parts. The first algorithm only takes into account the "average" inputs realizations, and, therefore, may not correctly capture the different time scales associated with the different actions, while the second takes into account the distribution of the inputs. While our optimization algorithms have been developed in the context of optimizing WWTPs' operations, we fully expect them to be applicable to a variety of complex control scenarios.

Break.

Session 4. 4:15 - 7:00 PM. Keh-Wei Lih AT&T Labs, Chair

Rare-event Analysis and Simulation for Random Elliptic Differential Equations.

Liu Jingchen, **Columbia University.**

In this talk, we consider an elliptic partial differential equation that admits a spatially varying random coefficient taking a lognormal form. This equation has been widely used to describe systems in various fields such as material science, electrostatics, and hydrology. Of interest is the asymptotic analysis and the efficient computation of some rare events associated to this random equation whose interpretation varies in different contexts. The central technique is the introduction of a change of measure that is not of the traditional exponential tilting form. The proposed algorithm can be proved to be efficient for the one-dimensional equation (ordinary differential equation) and it admits good empirical performances

Optimal Control of a Mean-Reverting Inventory and Optimal Production Management when Demand Depends on the Business Cycle.

Abel Cadenillas, **University of Alberta, Canada**, Peter Lakner, **and** Pinedo Michael, **NYU - Stern School of Business.**

We present two papers in which we apply the theory of stochastic control to solve management problems. In the first paper, motivated by empirical observations, we assume that the inventory level of a company follows a mean-reverting process. The objective of the management is to keep this inventory level as close as possible to a given target, so there is a running cost associated with the difference between the actual inventory level and the target. If inventory deviates too much from the target, management may perform an intervention in the form of either a purchase or a sale of the goods. There are fixed and proportional costs associated with each intervention. The management wants to find the optimal inventory levels at which interventions should be performed as well as the magnitudes of the interventions to minimize the total cost. We solve this problem by applying the theory of stochastic impulse control. In the second paper, we assume that the cumulative consumer demand for an item follows a Brownian motion, with both the drift and the variance parameters modulated by a continuous-time Markov chain that represents the regime of the economy. The management of the company would like to maintain the inventory level as close as possible to a target inventory level and would

also like to produce at a rate that is as close as possible to a target production rate. The company is penalized for deviations from the target levels, and the objective is to minimize the total discounted penalty costs. We consider two models. In the first model the management of the company knows the state of the economy, whereas in the second model the management does not know it. We solve both problems, and obtain the optimal production policy and the minimal total expected discounted cost. Furthermore, we compare the total expected discounted costs of the two models and determine the value of knowing the regime of the economy. We also solve the above problems in the case when the consumer demand rate follows a geometric Brownian motion modulated by the continuous-time Markov chain that represents the regime of the economy.

Impulsive control for continuous-time Markov Decision Processes.

Alexey Piunovskiy, **University of Liverpool and Franois Dufour**, **INRIA Bordeaux Sud-Ouest, France.**

We suggest a new mathematical model for continuous-time MDP, where along with the usual gradual control one can apply interventions, i.e. instantly change the state of the process. Any finite instant series of impulses is allowed. For the discounted objective, the dynamic programming approach is justified, and the sufficient class of control strategies is described. Those are stationary non-randomized ones with the possible interventions only immediately after natural jumps (and maybe at the initial time moment). For the constrained multiple-objective case, we develop the convex analytic approach: investigate the occupation measures, formulate and study the linear programs on the measure space. Stationary randomized strategies similar to those mentioned above, are sufficient for solving such constrained problems.

Stochastic Compositional Gradient: Optimizing Compositions of Stochastic Functions.

Mengdi Wang, **Princeton University.**

Classical stochastic gradient methods are well suited for minimizing expected values. However, they do not apply to the minimization of a nonlinear function involving expected values, i.e., $\min_x f(E_w[g_w(x)])$. We propose a stochastic compositional gradient descent (SCGD) method that updates using sample subgradients and a extrapolation step. For smooth problems, the SCGD converges at a rate of $O(k^{-2/7})$ in the general case and $O(k^{-4/5})$ in the strongly convex case.

Control of Non-Conventional Queueing Networks: A Parametric and Approximate Dynamic Programming Approach.

Emmanuel Fernandez, **Xiaoting Chen**, **University of Cincinnati.**

We investigate the control of non-conventional queueing networks, where multiple concurrent state transitions following non-exponential and general sojourn distributions are allowed. Two approximation schemes are discussed that produce an approximated Markovian model. We further propose to model the problem as an uncertain Markov decision processes (MDP), by considering the induced approximation error. A new simulation-based approach is investigated here, to give an overall optimal policy beyond the classic approach to such problems, e.g., a robust formulation. In particular, we approach the problem as one of finding a best overall control policy, via exploration and exploitation within a heuristic policy set. An approximate dynamic programming algorithm is then used, in connection with a parametric cost function, for efficiently learning and finding policies in this set. We show that the problem of finding the best control policy within this new policy set can be understood, equivalently, as one of finding the best set of parameters for one-stage cost function of the problem. Later, an integrated framework, denoted as extended actor-critic, is proposed to give a comprehensive treatment for this type of formulations. Results of a case study are also presented and discussed.

On Optimizing a constant production rate subject to inventory holding lost sales costs.

Melamed Benjamin, **Rutgers University**, **Junmin Shi**, **Georgia State University.**

In this talk we consider a single-product production - inventory system under continuous review with constant replenishment and compound Poisson demands subject to lost-sales. The system incurs carrying costs and lost-sale penalties, where the carrying cost is a linear function of the on-hand inventory and the lost-sales penalty is incurred per lost-sale occurrence and it is a function of the lost-sale size. We first derive an integro-differential equation for the expected cumulative cost until and including the first lost-sale occurrence. From this equation, we obtain a closed form expression for the time-average inventory cost, and provide an algorithm for a numerical computation of the optimal replenishment rate that minimizes the aforementioned time-average cost function. In particular, we consider two special cases of lost-sales penalty functions: constant penalty and loss-proportional penalty. We discuss simplifications for several parametric demand size distributions, such as constant, uniform and Gamma. For the special case of exponential demand sizes, we exhibit a closed form expression for the optimal replenishment rate and its corresponding cost. Finally, we illustrate the results with a numerical study.

7:00 - 9:00 PM Dinner. **Pay Tribute to Eugene A. Feinberg**, Isaac Sonin and Michael N. Katehakis, facilitators.

Organized by: Michael N. Katehakis (Chair)

Lei Lei Dimitris Metaxas and Joseph Mitchell (co-Chairs)

Local Committee: Cowan W., Gursoy K., Hunt D., Papayanopoulos L., Rose R., Lih K-W, and Smit L.C.
Conference Coordinator: Ms. Luz Kosar kosar@andromeda.rutgers.edu
Registration: By email to the conference coordinator.