Castle Laboratory

Department of Operations Research and Financial Engineering

Final Project Report:

Information Acquisition and Representation Methods for Real-Time Asset Management

Principal Investigator:

Warren B. Powell
Princeton University
Department of Operations Research and Financial Engineering
Princeton, NJ 08544
powell@princeton.edu

Prepared for:

Air Force Office of Scientific Research
Grant: FA9550-08-1-0195

February, 2011

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
21 APR 2011		**01-07-2007 to 30-11-2011**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Information Acquisition And Representation Methods For Real-Time Asset Management	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Princeton University,Department of Operations Research and Financial Engineering,Sherrerd Hall,Princeton,NJ, 08544**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**25**	

Contents

1. Summary of activities

The last three years have been exceptionally productive. Our research focused on two complementary themes: optimal learning, which addresses the efficient collection of information, and approximate dynamic programming, which is a modeling and algorithmic strategy for solving complex, sequential decision problems. These problems arise in the control of complex machinery, R&D portfolio optimization, materials science (sequential design of experiments), communications, and a wide range of resource allocation problems that arise in operations and logistics including mid-air refueling, spare parts management, emergency response, and robust allocation of fuel, medical supplies and food.

In the process of making advances in approximate dynamic programming, we found ourselves making contributions to an area that is proving to be critical to both lines of investigation: machine learning. In fact, we have come to realize that machine learning is starting to play a critical role in the advancement of our ability to solve complex stochastic programming problems, and it began to play an important role both in optimal learning and approximate dynamic programming.

We have found it useful to think of stochastic optimization problems in terms of three closely related mathematical problems. These include:

Stochastic search:

$$\max_{x \in \mathcal{X}} \mathbb{E} F(x,W) \tag{1}$$

Policy optimization

$$\max_{\pi \in \Pi} \mathbb{E}\left\{ \sum_{t=0}^{T} \gamma^t C(S_t, X^{\pi}(S_t)) \mid S_0 \right\} \tag{2}$$

Dynamic programming

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C(S_t, x) + \gamma \mathbb{E}\left\{ V_{t+1}(S^M(S_t, x, W_{t+1})) \mid S_t \right\} \right. \tag{3}$$

Here, we assume that x is a decision, which may be a multidimensional, and even high-dimensional, vector. W is a vector of random variables. $X^{\pi}(S_t)$ is a function (policy) that determines a decision x given the information in the state variable S_t. In all of the above, we assume that the expectation cannot be computed exactly, either because the vector W is too complex, or perhaps because we do not know the distribution, depending instead on observations from an exogenous process for sample realizations.

Equation (1) is the classical statement of a stochastic search problem, where we have to choose a deterministic set of parameters x to maximize an uncertain function. Our work in optimal learning focuses on problems where the function $F(x,W)$ is expensive to compute. For example, it might involve a field experiment (testing a new technology, moving a sensor, testing a policy for managing people) or running an expensive simulation. During our research, we encountered a variety of (1) which appears to be a new problem class, which we refer to as stochastic search with an observable state. This problem is written

$$\max_{x\in\mathcal{X}} \mathbb{E}F(S,x,W). \tag{1a}$$

In this problem, we first observe an exogenous state S, then we make a decision x, and finally we observe an exogenous outcome W that depends on S and x. Each time we make a decision, we do so in a different state S, which makes it hard to learn from past decisions, a feature that is fundamental to stochastic search algorithms.

Policy optimization (equation (2)), is mathematically equivalent to stochastic search (especially the form in equation (1a)), but the setting is typically different. A policy is some sort of rule for making decisions over time, and these come in many flavors.

The last problem class is dynamic programming, which is most familiar when written as Bellman's equation in (3). It is well know that this is a way of characterizing an optimal policy that solves (2), although this has never been viewed as an algorithmic strategy for stochastic search (equations (1) or (1a)).

It has long been recognized that statistical methods represent a powerful algorithmic strategy. Response surface methods (also known as metamodels) have long been recognized as a way of solving both stochastic search problems (1), and, since the 1990's, have been used as a powerful tool in the growing field of approximate dynamic programming for solving (3). However, the methods are often ad hoc since they depend on the "art" of feature selection (also known as basis functions). Convergence results (including our own contributions) tend to be limited to problems with special structure.

Our research has been progressing in parallel along three lines:

1. Machine learning – Both stochastic search and approximate dynamic programming depend on our ability to approximate either $\mathbb{E}F(x,W)$ (or $\mathbb{E}E(S,x,W)$), or the expected value function $\mathbb{E}\{V_{t+1}(S_{t+1})\,|\,S_t\}$. By far the most popular approximation strategy is to use a parametric representation which requires first manually identifying a set of features (or basis functions) which are typically denoted $\phi_f(S)$, $f\in\mathcal{F}$, which introduces the undesirable art of identifying features, which has grown into a side area of research. We started to pursue nonparametric methods, although classical techniques based on kernel regression do not scale to higher dimensions without assuming strong structural properties (although this remains an interesting area of research that we intend on pursuing). However, during the past three years, we made a significant advance

to a very general class of nonparametric methods known as Dirichlet process mixtures of generalized linear models (DP-GLM).

2. Optimal learning – There are a number of problems in stochastic search where the function $F(x,W)$ is expensive to measure, even for a single sample realization W. We developed a new search strategy called the *knowledge gradient* which we first discovered under our previous award, and which we have continued to develop in a significant way. Optimal learning is proving to be a powerful strategy for complex stochastic search problems, and we are just starting to investigate its use to solve the exploration vs. exploitation problem of approximate dynamic programming.
3. Approximate dynamic programming – We retain our original interest in solving sequential decision problems. These can sometimes be solved using policy optimization (equation (2)) as a form of stochastic search, but the most general strategy starts with Bellman's equation where we have to approximate the value function. In contrast with our previous research which focused on discrete resources (primarily motivated by problems in transportation and logistics), our work over the past three years has focused on states and actions that are both continuous and multidimensional, which have received relatively little attention in the stochastic optimization literature.

At this time, we have compiled theoretical and computational results that are starting to lend credence to the hope, long viewed as a kind of holy grail, that we might be able to develop *general purpose* solvers for the problems spanned by (1/1a), (2) and (3). While we doubt that a general purpose solver can outperform specialized solvers for specific problem class, there are parallels with the history of deterministic optimization where general purpose linear programming solvers replaced the specialized network codes, primal simplex codes and multicommodity codes that were popular in the 1980's. This is not to say that general purpose solvers can solve any integer or nonlinear programming problem, we can start to believe that we can significantly expand the range of stochastic control problems that can be solved using general purpose packages.

2. Technical advances

In this section, we summarize the research advances that we have made under the three general themes: machine learning, optimal learning and approximate dynamic programming.

2.1. Advances in machine learning

We began with the intent of using methods from machine learning to improve our ability to approximate value functions in ADP, and found ourselves instead making fundamental contributions to machine learning in the area of nonparametric statistics through joint research with Professor David Blei in computer science at Princeton. Lauren Hannah, funded by the AFOSR grant, began working with Prof. Blei and extended prior work on Dirichlet process mixtures to cover a broader class of problems that includes high-dimensional covariates which may be discrete, continuous or categorical. The ability to

handle high-dimensional covariates overcomes the central limitation of classical nonparametric statistics which uses kernel regression.

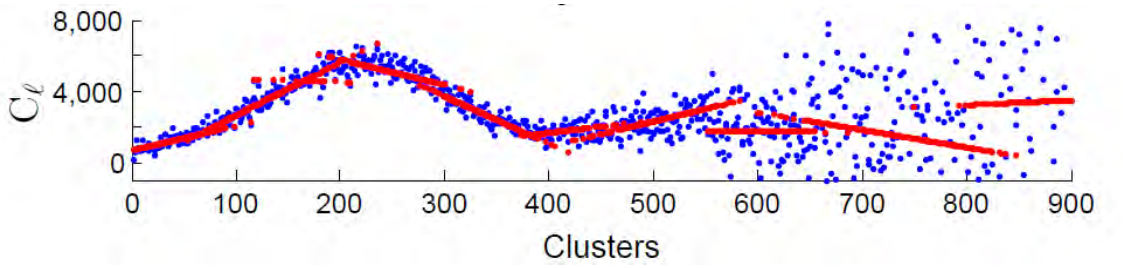

Figure 1 - Assigning data to clusters

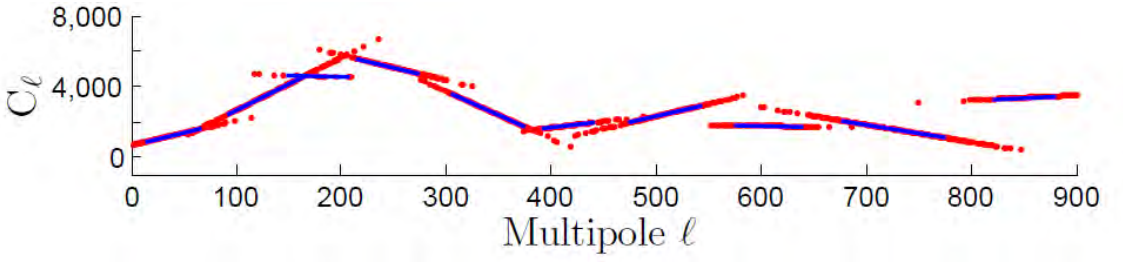

Figure 2-Fitting linear models to each cluster

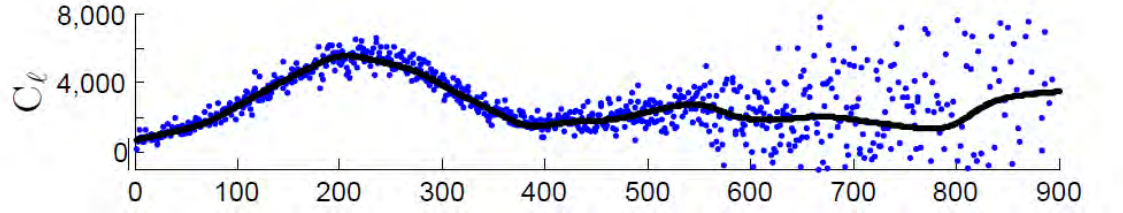

Figure 3-Smoothing the line using weighted estimates

The DP-GLM model is a Bayesian model where the response *y* and covariates *x* are characterized by a parameter vector $\theta_i = \left(\mu_i, \Sigma_i, \beta_i, \sigma_{\varepsilon,i}\right)$ where μ_i and Σ_i describes the mean and covariance matrix of the covariate vector $x_i \sim N\left(\mu_i, \Sigma_i\right)$ (explanatory variables) of the ith observation, while β_i is a vector of regression coefficients specifying the response. The parameter vector θ_i for the ith observation is assumed to belong probabilistically to one of a set of clusters. The probability it belongs to each cluster is given by a Dirichlet distribution, which is conjugate with the multinomial distribution describing the membership in a cluster. The response $y_i \mid x_i, \theta_i \sim N\left(\beta_{0,i} + \beta_{1:d,i}^T x_i, \sigma_{\varepsilon,i}^2\right)$ is assumed to be described by a linear regression, or any function in a broad class of generalized linear models. In a nutshell, DP-GLM can be viewed as a method that probabilistically classifies each data point into one of a series of clusters which adapt to the data.

Figure 1 illustrates the process of clustering observations. Figure 2 then shows the local linear fits to each cluster. Finally, figure 3 uses a weighting formula that estimates the probability that each data point is a member of each cluster to produce a smoothed fit.

In addition to the algorithm, Lauren Hannah was able to complete a very difficult proof of asymptotic unbiasedness, which means that this method offers the potential to approximate any problem. The paper can be downloaded by clicking on

> L. Hannah, D. Blei and W. B. Powell, "Dirichlet Process Mixtures of Generalized Linear Models," revised and resubmitted to *J. Machine Learning Research*. This paper is the central paper that introduces DP-GLM and provides the proof of asymptotic unbiasedness.

While this paper is under review (it has been revised and resubmitted), it was accepted for plenary presentation at the prestigious AISTATS conference:

> Hannah, L., D. Blei, W. B. Powell, "Dirichlet Process Mixtures of Generalized Linear Models," AISTATS, Italy, May, 2010. Selected for plenary presentation, which includes less than 10 percent of the submitted papers.

This strategy was recently extended to the problem in equation (1a) of stochastic search with an observable state variable. This problem arises in stochastic search problems where the solution depends on the "state of the world". If we have only one state of the world, we return to equation (1). If there are a small number of discrete states, we can solve this problem using an adaptation of classical stochastic search methods which perform updates (e.g. Robbins-Monroe stochastic gradient updates) which depend on the state of the world. This idea breaks down when the states are multidimensional and/or continuous. A draft of this paper can be downloaded from:

> Hannah, L., W. B. Powell, D. Blei, "Dirichlet Process Mixture Models for Stochastic Optimization with an Observable State Variable," in preparation for SIAM J. Optimization (should be submitted in May, 2010).

2.2. Optimal learning

The field of optimal learning (a name that we have introduced in an effort to help integrate the different communities that contribute to this problem) addresses the problem of collecting information when observations are expensive. We originally started working on this topic to solve the exploration vs. exploitation problem of approximate dynamic programming. As with our work on machine learning, this area of research took on a life of its own.

Our central contribution was the discovery that a "myopic policy" that we refer to as the knowledge gradient worked very well. The knowledge gradient is defined very simply. Let

y = Implementation decision (what we are going to do with the information)

K^n = State of knowledge (belief) about the value of different alternatives

$F(y,K)$ = The performance given decision y and knowledge K.

x^n = The choice of what to measure given K^n

The knowledge gradient is given by

$$\nu_{x^n}^{KG,n} = \mathbb{E}\left\{\max_y F\left(y, K^{n+1}(x^n)\right) - \max_y F(y,K^n)\right\},$$

which is effectively the economic value of measuring x^n. The KG policy is simply

$$x^n = \arg\max_x \nu_x^{KG,n}$$

Although the basic idea had been presented in a 1996 paper by Gupta and Miescke, we developed much more rigorously in

Frazier, P., W. B. Powell and S. Dayanik, "A Knowledge Gradient Policy for Sequential Information Collection," SIAM J. on Control and Optimization, Vol. 47, No. 5, pp. 2410-2439 (2008).

This is often dismissed as a myopic heuristic, but comparisons between this policy and one where decisions are optimized over a longer horizon suggest that the differences are negligible.

The original idea was developed for problems where we are trying to learn about discrete alternatives, and where learning something about one alternative teaches us nothing about another alternative (independent beliefs). A major practical breakthrough was the extension of this idea to the very important problem class of independent beliefs:

P. Frazier, W. B. Powell, S. Dayanik, "The Knowledge-Gradient Policy for Correlated Rewards," Informs Journal on Computing, Vol. 21, No. 4, pp. 585-598 (2009)

Most practical applications have correlated beliefs. Furthermore, this algorithm allows us to solve problems where the number of alternatives to measure may be much larger than our measurement budget.

The knowledge gradient is myopically optimal by construction; that is, it is the best measurement that you can make if you can make only one measurement. For offline problems, it is also asymptotically optimal, as both the papers above show. We also developed a general theory of asymptotic optimality that can be applied to other search policies:

P. Frazier and W. B. Powell, "Convergence to Global Optimality with Sequential Bayesian Sampling Policies" submitted to SIAM J. on Control and Optimization.

We often hear that many policies are asymptotically optimal (e.g. random search or round-robin), but the knowledge gradient is the only stationary policy that is both

myopically and asymptotically optimal, with the critical feature that it requires no tunable parameters.

The research above was performed in the context of offline learning problems. We recently adapted the idea to online learning problems, which are often referred to as multiarmed bandit problems. A special class of bandit problems can be solved optimally using a Gittins index policy, long viewed as a major breakthrough. However, computing Gittins indices is notoriously difficult, and the result cannot be generalized to problems with correlated beliefs.

> Ryzhov, I., W. B. Powell, P. I. Frazier, "The knowledge gradient algorithm for a general class of online learning problems", under review Operations Research (second revision).

This paper shows that the KG outperforms the best available approximation of the Gittins index on problems for which Gittins indices are optimal. However, the knowledge gradient can also handle finite horizon problems, as well as problems with correlated beliefs. Finally, this paper demonstrates that both offline and online problems can be solved using the same strategy (there is a trivial difference in the formulas) which is easily computable, and requires no tunable parameters.

What is perhaps the only limitation that we have been able to identify in the knowledge gradient is that some problems exhibit nonconcavity in the value of information. The value of one observation may be minimal, but 10 observations might be quite valuable. We can be led astray if we make measurement choices based on the value of a single measurement. The essential insight is that we only learn from a measurement when it is made with sufficient precision to change a decision. We overcome this limitation using a very simple, and easily computable, modification of the knowledge gradient that we are calling the KG(*) algorithm.

We have been extending the knowledge gradient to different problem classes. One involves learning about the edges in a graph. Consider the wide range of graph problems, and assume that we have imperfect information about the cost of an edge. We can use the knowledge gradient to determine which edge we should collect information about. This work is summarized in

> Ilya Ryzhov and W. B. Powell, "Information collection on a graph," Operations Research (to appear).

This paper means that we can quickly adapt the knowledge gradient policy for any offline problem to an online problem.

We are also nearing completion of an adaptation of KG to problems where we are measuring continuous parameters, as often arises when tuning the parameters of a physical device, experiment or the parameters of a simulation. The first step in this research is nearing completion and can be viewed at

> W. Scott, P. Frazier, W. B. Powell – "The Correlated Knowledge Gradient for Maximizing Expensive Continuous Functions with Noisy Observations using Gaussian Process Regression." In preparation (should be submitted May, 2010).

The challenge with continuous measurements is that the choice of measurement x is now a multidimensional continuous vector. As a result, solving $\arg\max_x \nu^{KG}(x)$ requires solving a nonlinear programming problem. We use an approximation of the knowledge gradient to derive analytical expressions for derivatives. $\nu^{KG}(x)$ is a nonconvex surface, depicted below.

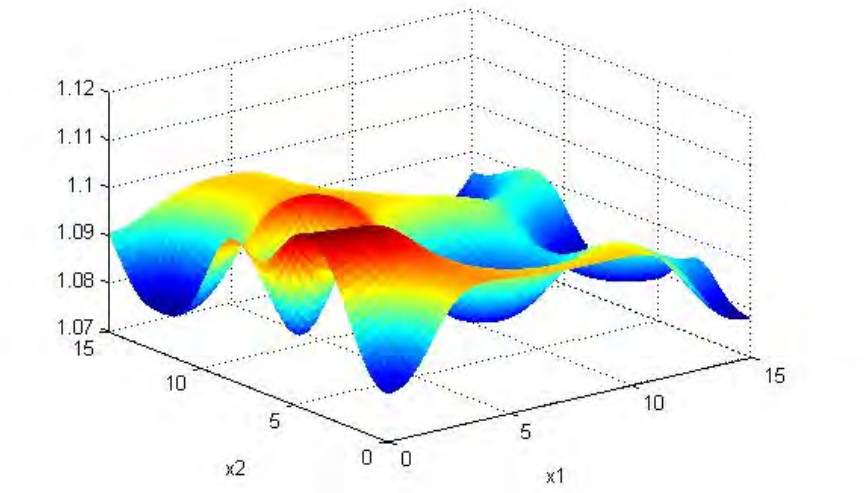

Figure 4 - Example of a 2-dimensional knowledge gradient surface

We have also been adapting the knowledge gradient to different types of beliefs. The three papers below adapt the knowledge gradient to problems with parametric beliefs (linear regression), beliefs based on weighted hierarchical estimates, and nonparametric beliefs.

D. Negoescu, P. Frazier and W. B. Powell, "The Knowledge Gradient Algorithm for Sequencing Experiments in Drug Discovery", Informs Journal on Computing (under revision). Received honorable mention in the Informs "Doing Good with Good OR."

Mes, M., P. I. Frazier and W. B. Powell, "Hierarchical Knowledge Gradient for Sequential Sampling," submitted to J. Machine Learning Research, November 19, 2009.

E. Barut, W. B. Powell, "Optimal Learning for Sequential Sampling with Non-Parametric Regression"

The paper on drug discovery made it possible to find the best molecular compound, out of 87,000 combinations, in under 200 trials. The work on hierarchical knowledge gradient is a simple form of nonparametric estimation, which makes it possible to optimize over very complex surfaces. The paper includes a convergence proof. The last paper uses classical kernel regression and also includes a convergence proof. This algorithm was used this past semester in several projects involving policy optimization, but at the moment it is limited to only a few continuous parameters.

Our next step is to see if we can adapt the knowledge gradient when the belief structure is represented using the DP-GLM model.

2.3. Approximate dynamic programming

After years of working on approximate dynamic programming for discrete resources, we shifted gears a few years ago to do convergence theory for ADP for problems with continuous, multidimensional states and actions. Virtually any ADP algorithm can handle complex states (this is the central goal of ADP), but most convergence proofs have been done in the reinforcement learning literature for problems where actions are discrete (or discretized). A popular strategy in this community, which avoids the explicit computation of the expectation (which is generally impossible) is to use the concept of Q-

learning, where instead of approximating the value of being in a state, $V(S)$, we estimate the value of a state action pair, denoted $Q(S,a)$, where a is a discrete action. Obviously estimating $Q(S,a)$ is harder than estimating $V(S)$, but if the action space is small, then it is not too much harder.

We are interested in problems where the action is a continuous vector x. In this setting, estimating $Q(S,x)$ is now dramatically harder than estimating $V(S)$ (throughout our discussion, we are using what the community refers to as "model-based" dynamic programming, where we assume we know the transition function).

We now face several technical challenges:

1. How do we solve for the vector x when there is an imbedded expectation?
2. How do we approximate the value function?
3. How do we solve the exploration vs. exploitation problem in high dimensions?
4. How do we perform statistical updating?

We solve the problem of the imbedded expectation by using the idea of the post-decision state, which is the value of a state, typically denoted S_t^x after a decision is made but before any new information has arrived, which means it is a deterministic function of the state S_t and action x_t. We developed this idea earlier and have demonstrated its effectiveness in a variety of transportation applications.

The last question represents a serious challenge when we use a particular algorithmic strategy that is variously called approximate value iteration, or TD(0) learning. This is the easiest strategy to implement computationally, since it means that we solve a sequence of deterministic optimization problems of the form

$$\max_{x\in\mathcal{X}} \left(C(S_t^n,x)+\bar{V}_t^x(S_t^x)\right).$$

This can typically be solved using a commercial solver for linear, nonlinear or integer programs. Approximate value iteration, however, requires updating of the general form

$$\bar{V}^n(S^n)=(1-\alpha_{n-1})\bar{V}^{n-1}(S^n)+\alpha_{n-1}\hat{v}^n$$

where $\hat{v}^n$ is new information about the value of being in state S^n. We found that when using approximate value iteration, considerable care has to be applied in the choice of stepsize formula. For this reason, we derived a new, optimal stepsize formula which appears to be the first optimal stepsize derived specifically for dynamic programs. The formula is presented below, along with a number of other insights about stepsizes:

Ryzhov, I., P. I. Frazier and W. B. Powell, "Stepsize Selection for Approximate Value Iteration and a New Optimal Stepsize Rule," submitted to J. Machine Learning Research, November 15, 2009.

We have used approximate policy iteration in most of our applications of approximate dynamic programming for the management of physical resources. In one special case, which arises when there are sequences of problems linked by a scalar variable as might arise in a storage application, we could prove convergence using approximate value iteration. This paper can be viewed at

> J. Nascimento, W. B. Powell, "An Optimal Approximate Dynamic Programming Algorithm for the Energy Dispatch Problem with Grid- Level Storage," under review at SIAM J. Control and Optimization.

We then undertook the problem of proving convergence for ADP algorithms designed specifically for this problem class. Our first paper assumes that we can exactly represent the value function (around the post-decision state) using basis functions (a parametric representation). We were una

> J. Ma and W. B. Powell, "Convergence Analysis of On-Policy LSPI for Multi-Dimensional Continuous State and Action-Space MDPs and Extension with Orthogonal Polynomial Approximation," under review at SIAM J. Control and Optimization.

A disappointment was that we had to resort to approximate policy iteration rather than approximate value iteration. Approximate policy iteration introduces an inner loop where we have to ensure that we do a "good enough" job of updating the value function. This was not needed in the previous reference with the scalar storage component. The last paper also required that we precisely know the basis functions, although it is shown that we can avoid this if we use orthogonal polynomials.

A key feature of this algorithm is that it is "on policy." This means that if we are in a state S^n and choose action x^n, the next state we visit is given by the transition function $S^{n+1} = S^M\left(S^n, x^n, W^{n+1}\right)$ where W^{n+1} is a Monte Carlo sample of the random information in *W*. This basic operation scales to very high dimensions (as we have found in our transportation work). But it means that the next state we visit is determined by our policy, which is generally not the correct policy. Most ADP/RL algorithms use off-policy sampling, where after optimizing the approximate decision function, an action is chosen at random to determine the next state to visit (we could also simply sample a state at random). Sampling an action at random is easy if there is a small number of discrete actions, but becomes meaningless when *x* is multidimensional (and especially if it is high dimensional). Off-policy sampling makes it easy to prove convergence with guarantees that states may be visited infinitely often, but computationally, it is completely impractical.

Our algorithm has three nice features: a) it does not require approximation of Q factors (around the state and action), b) it uses on-policy iteration, and c) it does not require an explicit exploration/exploitation strategy. The last feature arises (with some reasonable assumptions) because we only need to sample enough states to solve the identification problem for the parameters of the value function approximation.

The major limitation of this algorithm is that it requires that the value function be exactly represented by known basis functions, a condition that will never be satisfied in practice. For this reason, we turned next to studying theoretical convergence of an algorithm that approximates the value function using kernel regression, eliminating the need to know basis functions. This paper is nearing completion, and can be viewed by clicking on

> J. Ma and W. B. Powell, "Convergence Analysis of On-Policy LSPI for Multi-Dimensional Continuous State and Action-Space MDPs and Extension with Orthogonal Polynomial

Approximation," under review at IEEE Transactions on Automatic Control (likely submission May, 2010).

These two papers lay the theoretical foundation for provably convergent algorithms designed for continuous, multidimensional (possibly high dimensional) states and actions which depend on machine learning techniques to approximate the value function.

3. Selected applications

Energy

Drug discovery

Spare parts

Schneider

4. Research reports sponsored by AFOSR (2008-2010)

4.1. Journal articles

My papers strike a balance between theory and application. **Papers with substantial theoretical content are marked in bold.**

4.1.1. Under review

These papers are the best indication of recent research productivity.

1. **J. Ma and W. B. Powell, "Convergence Analysis of On-Policy LSPI for Multi-Dimensional Continuous State and Action-Space MDPs and Extension with Orthogonal Polynomial Approximation," under review at IEEE Transactions on Automatic Control.**
2. **Ryzhov, I., W. B. Powell, P. I. Frazier, "The knowledge gradient algorithm for a general class of online learning problems", under review at Operations Research (second round).**
3. Frazier, P. I., and W. B. Powell, "Paradoxes in Learning: The Marginal Value of Information and the Problem of Too Many Choices," submitted to Decision Analysis.
4. **J. Ma and W. B. Powell, "Convergence Analysis of On-Policy LSPI for Multi-Dimensional Continuous State and Action-Space MDPs and Extension with Orthogonal Polynomial Approximation," submitted to SIAM J. Control and Optimization.**
5. George, W. B. Powell, B. Bouzaiene-Ayari, J. Berger, A. Boukhtouta, "An Adaptive Learning Framework for Semi-Cooperative Multi-agent Cooperation," submitted to European Journal of Operations Research (this paper uses approximate dynamic programming to show that you can obtain near-optimal solutions even in a multiagent setting).
6. **Mes, M., P. I. Frazier and W. B. Powell, "Hierarchical Knowledge Gradient for Sequential Sampling," submitted to J. Machine Learning Research.**

7. **Ryzhov, P. I. Frazier and W. B. Powell, "Stepsize Selection for Approximate Value Iteration and a New Optimal Stepsize Rule," submitted to J. Machine Learning Research.**
8. **P. Frazier and W. B. Powell, "Convergence to Global Optimality with Sequential Bayesian Sampling Policies" submitted to SIAM J. on Control and Optimization.**
9. **J. Nascimento, W. B. Powell, "An Optimal Approximate Dynamic Programming Algorithm for the Energy Dispatch Problem with Grid- Level Storage," submitted to SIAM J. Control and Optimization (second round).**
10. D. Negoescu, P. Frazier and W. B. Powell, "The Knowledge Gradient Algorithm for Sequencing Experiments in Drug Discovery", Informs Journal on Computing.
11. **L. Hannah, D. Blei and W. B. Powell, "Dirichlet Process Mixtures of Generalized Linear Models," under review at J. Machine Learning Research.**
12. W. B. Powell, B. Bouzaiene-Ayari, J. Berger, A. Boukhtouta, A. George, "The Effect of Robust Decisions on the Cost of Uncertainty in Military Airlift Operations," submitted to ACM TOMACS.

4.1.2. Accepted

1. **Ilya Ryzhov and W. B. Powell, "Information collection on a graph," Operations Research (to appear).**
2. **L. Hannah and W. B. Powell, "Proof of Convergence for Evolutionary Policy Iteration under a Sampling Regime," IEEE Transactions on Automatic Control (to appear).**
3. Powell, W.B., "Merging AI and OR to Solve High-Dimensional Resource Allocation Problems using Approximate Dynamic Programming" Informs Journal on Computing, Vol. 22, No. 1, pp. 2-17 (2010).
4. L. Hannah, W. B. Powell, and J. Stewart, "One-Stage R&D Portfolio Optimization with an Application to Solid Oxide Fuel Cells," Energy Systems Journal, Vol. 1, No. 1, 2010.
5. **P. Frazier, W. B. Powell, S. Dayanik, "The Knowledge-Gradient Policy for Correlated Rewards," Informs Journal on Computing, Vol. 21, No. 4, pp. 585-598 (2009).**
6. Wu, Tongqiang, W.B. Powell and A. Whisman, "The Optimizing-Simulator: An Illustration using the Military Airlift Problem," ACM Transactions on Modeling and Simulation, Vol. 19, No. 3, Issue 14, pp. 1-31 (2009).
7. Simao, H. P., J. Day, A. George, T. Gifford, J. Nienow, W. B. Powell, "An Approximate Dynamic Programming Algorithm for Large-Scale Fleet Management: A Case Application," Transportation Science, Vol. 43, No. 2, pp. 178-197 (2009).
8. Wu, Tongqiang, W.B. Powell and A. Whisman, "The Optimizing-Simulator: An Illustration using the Military Airlift Problem," ACM Transactions on Modeling and Simulation, Vol. 19, No. 3, Issue 14, pp. 1-31 (2009).

9. Simao, H. P., J. Day, A. George, T. Gifford, J. Nienow, W. B. Powell, "An Approximate Dynamic Programming Algorithm for Large-Scale Fleet Management: A Case Application," Transportation Science, Vol. 43, No. 2, pp. 178-197 (2009).

10. Powell, W. B. "What you should know about approximate dynamic programming," Naval Research Logistics, Vol. 56, No. 3, pp. 239-249, 2009.

11. Simao, H. P. and W. B. Powell, "Approximate Dynamic Programming for Management of High Value Spare Parts", Journal of Manufacturing Technology Management Vol. 20, No. 2, pp. 147-160 (2009).

12. **Nascimento, J. and W. B. Powell, "An Optimal Approximate Dynamic Programming Algorithm for the Lagged Asset Acquisition Problem," Mathematics of Operations Research, Vol. 34, No. 1, pp. 210-237 (2009).**

13. George, A., W.B. Powell and S. Kulkarni, "Value Function Approximation Using Hierarchical Aggregation for Multiattribute Resource Management," Journal of Machine Learning Research, Vol. 9, pp. 2079-2111 (2008).

14. **S. Dayanik, W. Powell, and K. Yamazaki, "Index policies for discounted bandit problems with availability constraints," Advances in Applied Probability, Vol. 40, No. 2, pp. 377-400 (2008).**

15. **Frazier, P., W. B. Powell and S. Dayanik, "A Knowledge Gradient Policy for Sequential Information Collection," SIAM J. on Control and Optimization, Vol. 47, No. 5, pp. 2410-2439 (2008).**

16. Cheung, R. K.-M., N. Shi, W. B. Powell, and H. P. Simao, "An Attribute-Decision Model for Cross-Border Drayage Problem," Transportation Research E: Logistics and Transportation Review, Volume 44, No. 2, pp. 217-234 (2008).

4.2. Refereed book chapters and conference proceedings

4.2.1. To appear

1. Powell, W. B., "The Knowledge Gradient for Optimal Learning," Encyclopedia for Operations Research and Management Science (to appear).

2. Ryzhov, I. O., P. I. Frazier, W. B. Powell, "On the Robustness of a One-Period Look-ahead Strategy for Multi-armed Bandit Problems," International Conference on Computer Science, Amsterdam, May, 2010.

3. Powell, W. B., "Approximate Dynamic Programming I: Modeling," Encyclopedia of Operations Research and Management Science, John Wiley and Sons (to appear).

4. Powell, W. B., "Approximate Dynamic Programming II: Algorithms," Encyclopedia of Operations Research and Management Science, John Wiley and Sons (to appear)

4.2.2. Appeared

1. **Hannah, L., D. Blei, W. B. Powell, "Dirichlet Process Mixtures of Generalized Linear Models," AISTATS, Italy, May, 2010.**

2. Ryzhov, I., W. B. Powell, "A Monte-Carlo Knowledge Gradient Method for Learning Abatement Potential of Emissions Reduction Technologies," Winter

Simulation Conference, 2009. M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, eds, 2009, pp. 1492-1502.

3. Frazier, P., W. B. Powell, H. P. Simao, "Simulation Model Calibration with Correlated Knowledge-Gradients," Winter Simulation Conference, M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, eds, 2009, pp. 339-353.

4. **Ma, J. and W. B. Powell, "A convergent recursive least squares policy iteration algorithm for multi-dimensional Markov decision process with continuous state and action spaces," IEEE Conference on Approximate Dynamic Programming and Reinforcement Learning (part of IEEE Symposium on Computational Intelligence), March, 2009.**

5. Ryzhov, I. and W. B. Powell, "The Knowledge Gradient Algorithm For Online Subset Selection," IEEE Conference on Approximate Dynamic Programming and Reinforcement Learning (part of IEEE Symposium on Computational Intelligence), March, 2009.

6. P. Frazier, W. B. Powell, S. Dayanik and P. Kantor, "Approximate Dynamic Programming in Knowledge Discovery for Rapid Response," HICSS Conference, 2009.

7. S Dayanik, W. B. Powell and K. Yamazaki "An Asymptotically Optimal Strategy in Sequential Change Detection and Identification Applied to Problems in Biosurveillance" Proceedings of the 3rd INFORMS Workshop on Data Mining and Health Informatics, (J. Li, D. Aleman, R. Sikora, eds.), 2008.

8. **Frazier, P. and W. B. Powell, "The knowledge gradient stopping rule for ranking and selection," Proceedings of the Winter Simulation Conference, December 2008.**

9. H. P. Simao and W. B. Powell, "Approximate Dynamic Programming for Managing High Value Spare Parts," Journal of Manufacturing Technology Management (to appear). Also recipient of Best Paper Prize at 2008 ICPR Americas Conference.

10. Powell, W. B., "Approximate Dynamic Programming: Lessons from the field," Invited tutorial, Proceedings of the 40th Conference on Winter Simulation, pp. 205-214, 2008.

11. Powell, W. B. and P. Frazier, "Optimal Learning," TutORials in Operations Research, Chapter 10, pp. 213-246, Informs (2008).

4.3. Books

1. Wiley has approved submission of a second edition of my book: *Approximate Dynamic Programming: Solving the curses of dimensionality.* This edition will include a number of advances from the last three years of research. I view this book as an important educational device.

2. *Wiley has given us a contract for a new book to be called Óptimal Learning, which is being written jointly with Ilya Ryzhov. We have written 200 pages, and anticipate submitting a manuscript in the fall of 2011.*

4.4. Doctoral dissertations

The following doctoral dissertations were completed over the last three years.

Peter Frazier, 2009 – "Knowledge Gradient Methods for Statistical Learning," First position: Cornell University, Department of Operations Research and Information Engineering.

Kazutoshi Yamazaki, 2009 – "Essays on Sequential Analysis: Multi-Armed Bandit with Availability Constraints and Sequential Change Detection and Identification," First position: Osaka University, Center for the Study of Finance and Insurance.

A third, by Lauren Hannah, will be finished this summer. Lauren was awarded a competitive fellowship at Duke University which is generally used to attract women and minorities into faculty positions at Duke.

5. Personnel supported

Faculty:

- Professor Warren B. Powell

Professional staff:

- Dr. Hugo Simao

Graduate students:

- Lauren Hannah (5^{th} year) – Ph.D.
- Ilya Rhyzov (4^{th} year) – Ph.D.
- Warren Scott (3^{rd} year) – Ph.D.
- Jae Ho Kim (3^{rd} year) – Ph.D.
- Emre Barut (2^{nd} year) – Ph.D.

6. Honors and awards

Winner, Donald H. Wagner Prize for Excellence in Operations Research Practice, Fall, 2009. This award was given for an industrial application of approximate dynamic programming, which was funded over the years by my AFOSR research. The Wagner prize is specifically designed to recognize contributions to methodology arising from practice.

Honorable mention – Doing Good with Good OR, student paper competition run by Informs, Fall, 2009.

Best paper prize at ICPR Americas conference, June, 2008, "Approximate Dynamic Programming for Managing High Value Spare Parts." (with H. P. Simao)

7. Interactions/transitions (2008-2010).

7.1. Participation/presentations at meetings, conferences, etc.

7.1.1. Invited talks:

1. "Optimal Learning," North Carolina State University, Raleigh, NC, April, 2010.
2. "Optimal Learning for Homeland Security," CCICADA Workshop, Morgan State, Baltimore, Md., March 7, 2010.
3. "Opportunities for Machine Learning in Stochastic Optimization, with Applications in Energy Resource Planning," Seminar series in computational sustainability, Cornell University, Department of Computer Science, March 5, 2010.
4. "Approximate Dynamic Programming for Energy Resource Management," Invited presentation for mini-symposia at Neural Information Processing Society, Vancouver, December 10, 2009.
5. "Solving High-Dimensional Stochastic Optimization Problems using Approximate Dynamic Programming," Princeton Program for Applied and Computational Mathematics seminar series, Princeton University, November 23, 2009.
6. "Approximate Dynamic Programming for Very Large-Scale Graphs," AFOSR Workshop on Network Mathematics, Computing and Applications, Harvard University, November 18, 2009.
7. "Approximate Dynamic Programming for High-Dimensional Resource Allocation Problems," Lehigh University, Department of Industrial and Systems Engineering, Nov. 13, 2009.
8. "Optimal Learning," School of Industrial and Systems Engineering, Georgia Institute of Technology, Nov 3, 2009.
9. "Research in Energy Systems Design and Control," Princeton Environmental Institute, October 2, 2009.
10. "Approximate Dynamic Programming for Freight Transportation," Norfolk Southern Railroad, August 21, 2009.
11. "Optimal Learning," IBM T.J. Watson Research Center, September 28, 2009.
12. "Optimal Learning: Efficient Information Collection for the Department of Homeland Security," Rutgers University, August 12, 2009.
13. "Approximate Dynamic Programming for High-Dimensional Resource Allocation Problems," Plenary speaker, IEEE International Conference on Automation and Logistics, Shenyang, China, August 6, 2009.
14. "Approximate Dynamic Programming for High-Dimensional Applications," Invited plenary speaker, Multidisciplinary Symposium on Reinforcement Learning (MSRL), McGill University, Montreal, June, 2009.

15. "Approximate Dynamic Programming for High-Dimensional Problems in Energy Modeling," Cornell Workshop on Computational Sustainability, Cornell University, June, 2009.
16. "Tutorial: Optimal Learning," Dagstuhl workshop on Sampling-Based Optimization in the Presence of Uncertainty, Dagstuhl, Germany, April, 2009.
17. "Approximate Dynamic Programming: Solving the curses of dimensionality," Cornell University, April 15, 2009.
18. "Optimal Learning," Cornell University, April 14, 2009.
19. "Optimal Learning using the Knowledge Gradient Policy," Rutgers University, March 23, 2009.
20. "Optimal Learning," London School of Economics, February 6, 2009.
21. "Approximate Dynamic Programming: Solving the curses of dimensionality," University of Nottingham (England), February 4, 2009.
22. "Approximate Dynamic Programming: Solving the curses of dimensionality," University of Lancaster (England), February 3, 2009.
23. "Optimal Learning," University of Lancaster (England), February 2, 2009.
24. Invited tutorial: "Approximate Dynamic Programming: Making Simulations Intelligent," Winter Simulation Conference, Miami, December, 2008.
25. "Optimal Learning and Change Detection," Workshop on Homeland Security, Princeton University, December 5, 2008.
26. "SMART: A Stochastic Multiscale Energy Policy Model using Approximate Dynamic Programming," Department of Energy, Washington, D.C., December 1, 2008.
27. "SMART: A Stochastic Multiscale Model for Energy Policy Model," 2nd Annual Western Region Energy Workshop, organized by Lawrence Livermore National Laboratories, November 4, 2008.
28. "From Transportation to Energy: A History of CASTLE Laboratory," 2nd Annual Western Region Energy Workshop, organized by Lawrence Livermore National Laboratories, November 3, 2008.
29. Tutorial for Informs Computing Society: "Approximate Dynamic Programming" Informs Annual Meeting, Washington D.C., 2008.
30. Tutorial: "Optimal Learning" Informs Annual Meeting, Washington D.C., 2008 (with Peter Frazier)
31. "Approximate Dynamic Programming for High-Dimensional Problems," Duke University, September 17, 2008.
32. "Optimal Learning for Nuclear Detection," Rutgers University, DyDAn Center, September 15, 2008.
33. "A Multiscale Energy Policy Model," Western Region Energy Workshop, Berkeley, CA, September 11, 2008.
34. "Approximate Dynamic Programming: Solving the Curses of Dimensionality," Invited plenary speaker, ICPR Americas, Sao Paulo, Brazil, June 6, 2008.

35. "The Optimizing-Simulator for Capturing Real-World Military Operations," Air Mobility Command, Scott AFB, May 27, 2008.
36. "Approximate Dynamic Programming: Solving the Curses of Dimensionality," Invited plenary speaker, CIRRELT Workshop, Quebec City, May, 2008.
37. "Information collection and learning for nuclear detection," Rutgers University, April, 2008.
38. "Approximate Dynamic Programming for High-Dimensional Problems," Boston University, February 29, 2008.

7.1.2. Conference presentations with refereed papers/abstracts:

1. "On the Robustness of a One-Period Look-Ahead Policy for Multiarmed Bandit Problems," (with I. Ryzhov, P. Frazier), International Workshop on Computational Stochastics, Netherlands, June 1, 2010.
2. "A Monte Carlo Knowledge Gradient Method for Learning Abatement Potential of Emissions Reduction Technologies," Winter Simulation Conference, Austin, TX, December 14, 2009 (with I. Ryzhov).
3. "Simulation Optimization with Correlated Knowledge Gradient," Winter Simulation Conference, Austin, TX, December 14, 2009 (with P. Frazier and H. P. Simao)
4. "A Monte-Carlo Knowledge Gradient Method For Learning Abatement Potential Of Emissions Reduction," Winter Simulation Conference, Houston, 2009 (with I. Ryzhov).
5. "Simulation Model Calibration with Correlated Knowledge Gradients," Winter Simulation Conference, Houston, 2009 (with P. Frazier)
6. "A convergent recursive least squares policy iteration algorithm for multi-dimensional Markov decision process with continuous state and action spaces", IEEE Conference on Approximate Dynamic Programming and Reinforcement Learning, Nashville, March 31, 2009 (with Jun Ma)
7. "The Knowledge Gradient Algorithm For Online Subset Selection", IEEE Conference on Approximate Dynamic Programming and Reinforcement Learning, Nashville, March 31, 2009 (with Ilya Ryzhov).
8. "The Knowledge Gradient Stopping Rule for Ranking and Selection," Winter Simulation Conference, Miami, December, 2008 (with P. Frazier).
9. "Locomotive Optimization for Norfolk Southern Railroad Using Approximate Dynamic Programming," Informs Annual Meeting, Washington D.C., 2008 (with Belgacem Bouzaiene-Ayari, Clark Cheng, Ricardo Fiorillo)
10. "Monte Carlo Evolutionary Policy Iteration with Applications to Energy R&D Portfolio Optimization," Informs Annual Meeting, Washington D.C., 2008 (with Lauren Hannah and Jeffrey Stewart)
11. "A Dynamic Energy Resource Modeling System," Informs Annual Meeting, Washington D.C., 2008 (with Abraham George, Alan Lamont and Jeffrey Stewart).
12. "One-Stage R&D Portfolio Optimization with an Application to Solid Oxide Fuel Cells," Informs Annual Meeting, Washington D.C., 2008 (with Lauren Hannah and Jeff Stewart).

13. “Optimal Control of Disease Decisions in Controlled Ovarian Hyperstimulation,” Informs Annual Meeting, Washington D.C., 2008 (with Miao He and Lei Zhao)
14. “Asymptotic Theory of Sequential Change Detection and Identification,” Informs Annual Meeting, Washington D.C., 2008 (with Kazutoshi Yamazaki and Savas Dayanik).
15. “General Asymptotic Theory of Sequential Change Detection and Identification,” NIPS, 2008.

7.1.3. Other conference presentations:

1. “Approximate Dynamic Programming for Management of High Value Spare Parts,” Informs Annual Meeting, San Diego, CA, October, 2009. (with H. Simao)
2. “Regression with a Dirichlet Process-Generalized Linear Mixture Models,” Informs Annual Meeting, San Diego, CA, October, 2009. (with L. Hannah and D. Blei)
3. “Simulation Calibration with Correlated Knowledge Gradients,” Informs Annual Meeting, San Diego, CA, October, 2009. (with Peter Frazier and H. Simao)
4. “The Correlated Knowledge Gradient for Continuous Decision Variables,” Informs Annual Meeting, San Diego, CA, October, 2009. (with W. Scott and P. Frazier)
5. “Knowledge Gradients with Monte Carlo Simulation in Online Learning Problems,” Informs Annual Meeting, San Diego, CA, October, 2009. (with I. Ryzhov)
6. “Energy Policy Conditional Optimization using Dirichlet Process-Generalized Linear Model Mixture,” Informs Annual Meeting, San Diego, CA, October, 2009. (with L. Hannah)
7. “SMART: Stochastic, Multiscale Energy Policy Model,” Informs Annual Meeting, San Diego, CA, October, 2009. (with A. George, A. Lamont, J. Stewart)
8. “Optimal Control of Wind Storage Process with Continuous States and Actions with Advance Commitments,” Informs Annual Meeting, San Diego, CA, October, 2009. (with J. Kim)
9. “Hierarchical Knowledge-Gradient Policy for Sequential Sampling,” Informs Annual Meeting, San Diego, CA, October, 2009. (with Martijn Mes)
10. “Convergent Least Squares Policy Iteration Algorithm for High Dimensional Markov Decision Processes,” Informs Annual Meeting, San Diego, CA, October, 2009. (with J. Ma)
11. “Optimal Learning on a Graph,” Informs Annual Meeting, San Diego, CA, October, 2009. (with I. Ryzhov)
12. “SMART: A Stochastic Multiscale Energy Policy Model using Approximate Dynamic Programming.” Power Systems Modeling Conference, University of Florida, Gainesville, April, 2009 (with Abraham George, Jeffrey Stewart and Alan Lamont).
13. “One Stage R&D Portfolio Optimization with an Application to Solid Oxide Fuel Cells,” Power Systems Modeling Conference, University of Florida, Gainesville, April, 2009 (with Lauren Hannah).
14. “Convergent Approximate Dynamic Programming Algorithm for Continuous State and Action Spaces,” Informs Computing Society, Charleston, SC, January, 2009 (with Jun Ma).

15. “Approximate Dynamic Programming for Management of High-Value Spare Parts,” Informs Computing Society, Charleston, SC, January, 2009 (with Hugo Simao).
16. “The Knowledge Gradient Algorithm for Sequential Information Collection,” Informs Computing Society, Charleston, SC, January, 2009 (with Ilya Ryzhov and Peter Frazier).
17. “Optimal Control of Dosage Decisions in Controlled Ovarian Hyperstimulation,” Informs Annual Meeting, Washington D.C., 2008 (with M. He and L. Zhao).
18. “A Dynamic Energy Resource Modeling System,” “One Stage R&D Portfolio Optimization with an Application to Solid Oxide Fuel Cells”, Informs Annual Meeting, Washington D.C., 2008 (with A. George, A. Lamont and J. Stewart)
19. “One Stage R&D Portfolio Optimization with an Application to Solid Oxide Fuel Cells”, Informs Annual Meeting, Washington D.C., 2008 (with L. Hannah and J. Stewart)
20. “Asymptotic Theory of Sequential Change Detection and Identification” Informs Annual Meeting, Washington D.C., 2008 (with Kazutoshi Yamazaki and Savas Dayanik).
21. “Asymptotic Theory of Sequential Change Detection and Identification” Informs Annual Meeting, Washington D.C., 2008 (with Kazutoshi Yamazaki and Savas Dayanik).
22. “Monte Carlo Evolutionary Policy Iteration with Applications to Energy R&D Portfolio Optimization,” Informs Annual Meeting, Washington D.C., 2008 (with L. Hannah).
23. “Information Collection With A Physical State,” Informs Annual Meeting, Washington D.C., 2008 (with Ilya Ryzhov)
24. “Knowledge Gradient for Bandit Problems,” Informs Annual Meeting, Washington D.C., 2008 (with Ilya Ryzhov).
25. “Locomotive Optimization for Norfolk Southern using Approximate Dynamic Programming,” Informs Annual Meeting, Washington D.C., 2008 (with B. Bouzaiene-Ayari, C. Cheng, R. Fiorillo, J. Chang)
26. “Optimal Learning for the Newsvendor Problem,” Informs Annual Meeting, Washington D.C., 2008 (with Diana Negoescu and Peter Frazier)
27. “Convergence of Sequential Sampling Policies for Bayesian Information Collection Problems,” Informs Annual Meeting, Washington D.C., 2008 (with Peter Frazier)
28. “The Knowledge-Gradient Policy for Ranking and Selection with Correlated Normal Beliefs,” Informs Annual Meeting, Washington D.C., 2008 (with Peter Frazier)
29. “Approximate Dynamic Programming for the Single Machine Scheduling Problem,” ICPR Americas, Sao Paulo, Brazil, June, 2008 (with Debora Ronconi).
30. “Approximate Dynamic Programming for the Management of High Value Spare Parts,” ICPR Americas, Sao Paulo, Brazil, June, 2008 (with Hugo Simao).

7.2. Consultative and advisory functions

Presentation: "Approximate Dynamic Programming for Very Large-Scale Graphs," AFOSR Workshop on Network Mathematics, Computing and Applications, Harvard University, November 18, 2009. Co-organized by Bruce Suter at AFRL.

Interaction with Bob Wright discussing his use of novel ADP algorithms.

7.3. Transitions

Our transitions have occurred along three lines:

- Direct implementation of ideas through projects with the corporate partners of CASTLE Lab. This is the major path by which we test our ideas in the field. Industrial projects during 2008-2010 included work with Schneider National (one of the three largest truckload motor carriers), Netjets (largest fractional jet operator), Norfolk Southern Railroad (one of four class I railroads in the U.S.), and Embraer (major manufacturer of regional jets).

- Posting software on the internet. This summer, we will be posting two important pieces of software: 1) The knowledge gradient calculator, which allows people to experiment with different learning policies, and 2) the DP-GLM machine learning software.

- Licensing of software through local consulting firms for use in systems for their clients. CASTLE Lab has a relationship with Princeton Consultants, Inc. (www.princeton.com) which implements optimization and simulation models in transportation and logistics.

Specific transitions to the industrial partners of CASTLE Lab over the last three years include:

1. Transition: Optimizing simulator for fleet planning at Schneider National. We have calibrated a system that models the flows of approximately 5,000 drivers of different types. Schneider is interested in knowing what types of drivers are most valuable to the fleet (similar to AMC asking which aircraft types are most valuable). It is almost impossible to answer this question using "what if" analyses. Our logic produces, from one run, estimates of the gradients with respect to each type of driver. This project won the Wagner Prize from Informs in 2009.

   Recipient: Schneider National, the nation's largest truckload motor carriers.

2. Transition: Operational, tactical and strategic planning of locomotives. This system uses the optimizing simulator concept, and in particular makes heavy use of techniques for modeling incomplete information through low dimensional patterns. The system was recently approved for production at Norfolk Southern Railroad, making it the first successful production optimization model developed for operational use in North America.

Recipient:

Norfolk Southern Railroad, which uses the system both for strategic planning of the fleet size, and short-term tactical forecasting of surpluses and deficits.

3. Transition: We developed a system for optimizing high-value spare parts. This problem involves designing inventory policies for parts where the inventories are often zero (only a few locations will have even a single spare). We have to design policies for hundreds of spare parts, so that the aggregate inventory cost is below a certain level, and where we achieve specific targets on aggregate service.

   Recipient: Embraer