<div align="center">

09181 Abstracts Collection
Sampling-based Optimization in the Presence of Uncertainty
— Dagstuhl Seminar —

</div>

<div align="center">

Jürgen Branke[1], Barry Nelson[2], Warren Powell[3] and Thomas J. Santner[4]

[1] University of Warwick, UK
`juergen.branke@wbs.ac.uk`
[2] NW University - Evanston, US
`nelsonb@northwestern.edu`
[3] Princeton University, US
`powell@princeton.edu`
[4] Ohio State University, US
`tjs@stat.ohio-state.edu`

</div>

Abstract. From April 26 to April 30, the Dagstuhl Seminar 09181 "Sampling-based Optimization in the Presence of Uncertainty" was held in Schloss Dagstuhl – Leibniz Center for Informatics. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of working group results and ideas are put together in this paper.

Keywords. Optimal learning, optimization in the presence of uncertainty, simulation optimization, sequential experimental design, ranking and selection, random search, stochastic approximation, approximate dynamic programming

09181 Executive Summary – Sampling-based Optimization in the Presence of Uncertainty

This Dagstuhl seminar brought together researchers from statistical ranking and selection; experimental design and response-surface modeling; stochastic programming; approximate dynamic programming; optimal learning; and the design and analysis of computer experiments with the goal of attaining a much better mutual understanding of the commonalities and differences of the various approaches to sampling-based optimization, and to take first steps toward an overarching theory, encompassing many of the topics above.

Keywords: Optimal learning, optimization in the presence of uncertainty, simulation optimization, sequential experimental design, ranking and selection, random search, stochastic approximation, approximate dynamic programming

Joint work of: Branke, Jürgen; Nelson, Barry L.; Powell, Warren Buckler; Santner, Thomas J.

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2009/2116

Sequential parameter optimization

Thomas Bartz-Beielstein (FH Köln, DE)

In the experimental analysis of algorithms, two issues are still not sufficiently treated. Firstly, the performance of algorithms depends on their parameterizations— and of the parameterizations of the problem instances. However, these dependencies can be seen as means for understanding algorithm's behavior. Secondly, the non-determinism of evolutionary and other metaheuristic methods renders results distributions, not numbers.

We provide a comprehensive, effective and very efficient methodology for the design and experimental analysis of algorithms. We rely on modern statistical techniques for tuning and understanding algorithms from an experimental perspective. Therefore, we make use of the sequential parameter optimization (SPO) method that has been successfully applied as a tuning procedure to numerous heuristics for practical and theoretical optimization problems.

Keywords: Optimization

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/2115

Learning and Anticipation in Online Dynamic Optimization with Evolutionary Algorithms

Peter Bosman (CWI - Amsterdam, NL)

The focus of this talk is on how to design evolutionary algorithms (EAs) for solving stochastic dynamic optimization problems online, i.e. as time goes by. For a proper design, the EA must not only be capable of tracking shifting optima, it must also take into account the future consequences of the evolved decisions or actions. In this talk we discuss how to properly tackle stochasticity in this context. We point out how this naturally leads to evolving strategies rather than explicit decisions. This approach is illustrated with a running example and also applied to inventory management problems, an important real-world application area in logistics. Our results show, as a proof of principle, the feasibility and benefits of the approach.

Keywords: Dynamic optimization, online optimization, evolutionary algorithms

Aggregation and the sensitivity based optimization

Xi-Ren Cao (The Hong Kong Univ. of Science & Technology, CN)

We first review the aggregation techniques in sample path based learning and optimization of Markov decision processes (MDPs), including state aggregation, time aggregation and event-based aggregation. We then discussed a simple and intuitive clear explanation/derivation of the performance optimization theory of MDPs, i.e., the sensitivity-based approach, and based on it, we explain why and when the aggregation techniques lead to accurate optimal solutions.

Keywords: Aggregation, MDPs, Event-based optimization, Time aggregation

Ranking and Selection Tutorial

Stephen E. Chick (INSEAD - Fontainebleau, FR)

We review a number of different approaches to the problem of ranking and selection. Ranking and selection deals with finding the best of a finite set of alternatives, where best is determined with respect to the mean, and the mean is to be inferred from statistical sampling. The questions include which alternatives to sample and and how many times, whether to stop sampling or to continue with one or more rounds of sampling, and which alternative to select as best when the alternatives stops.

Keywords: Ranking and selection, indifference zone, OCBA, Bayesian sequential sampling

Probabilistic Models for Fast Learning in Control

Marc Deisenroth (Cambridge University, GB)

Learning from experience is a key ingredient in the behavior of intelligent beings and holds great potential for artificial systems. Humans and animals use experience to learn complicated tasks relatively quickly, that is, they do not require many trials to succeed. In contrast, artificial learners cannot generally replicate this speed of learning. To speed up artificial learning, we borrow two key ingredients that make biological learning so successful: the ability to generalize and the explicit incorporation of uncertainty into the decision-making process. We explicitly use probabilistic Gaussian process models for predictions to account for both key ingredients.

We successfully apply our learning algorithm to control problems with continuous state and action spaces.

Our algorithm learns to solve complicated tasks, such as the cart-pole swing up, the Pendubot, or the cart-double pendulum swing up, in a couple of trials.

Keywords: Gaussian processes, Bayesian inference, reinforcement learning, system identification, optimal control

Joint work of: Deisenroth, Marc Peter; Rasmussen, Carl Edward

Ranking and Selecton of Many Alternatives using Knowledge Gradients and Correlated Bayesian Beliefs

Peter Frazier (Princeton University, US)

We consider a Bayesian ranking and selection problem with normal rewards and a correlated normal prior. Because this formulation of the ranking and selection problem models dependence in the belief between alternatives, algorithms within it may perform efficiently even when the number of alternatives is very large. We propose a fully sequential policy called the correlated knowledge-gradient policy, which is provably optimal in some special cases and has bounded suboptimality in all others.

Keywords: Ranking and selection, bayesian global optimization, expected improvement, value of information

Joint work of: Frazier, Peter; Powell, Warren; Dayanik, Savas

Full Paper:
 http://www.princeton.edu/∼pfrazier/pub/CorrelatedKG.pdf

Stochastic Gradient Estimation: An Overview

Michael Fu (University of Maryland - College Park, US)

I provide a tutorial overview of methods for estimating gradients from stochastic simulation. Such estimators can be used for simulation-based optimization or sensitivity analysis. The main approaches described are finite differences, perturbation analysis, the likelihood ratio/score function method, and weak derivatives.

Keywords: Stochastic gradient estimation, perturbation analysis, likelihood ratio method, score function, weak derivatives

See also: M.C. Fu, "What You Should Know About Simulation and Derivatives" (Cover Story), *Naval Research Logistics*, Vol.55, No.8, 723–736, 2008. M.C. Fu, "Gradient Estimation," Chapter 19 in *Handbooks in Operations Research and Management Science: Simulation*, S.G. Henderson and B.L. Nelson, editors, Elsevier, 575–616, 2006.

Some Problems in Need of Sampling-Based Optimization Under Uncertainty Methods

Genetha Gray (Sandia Nat. Labs - Livermore, US)

In this talk, I will review some of the current problems that are in need of more efficient methods of sampling in the presence of uncertainty. These problems all require the results of computationally expensive black box functions. The goals include general sensitivity analysis, calibration (with real data that itself contains uncertainties), and validation. Although I will discuss some of the methods that have been tried on these problems, the real purpose of this talk is to inspire the audience to think of methods for sampling in the presence of uncertainty within the limitations inherent in a set of real problems from the engineering discipline.

Design and Analysis of Experiments: Overview

Jack P. C. Kleijnen (Tilburg University, NL)

This tutorial gives an overview of the design and analysis of experiments with optimization algorithms. It covers classic designs and their corresponding (meta)models; namely, Resolution-III designs (including fractional factorial two-level designs) for first-order polynomial models, resolution-IV and resolution-V designs for two-factor interactions, and designs (including central composite designs) for second-degree polynomials.

The tutorial also reviews factor screening in experiments with very many factors, focusing on the sequential bifurcation method. Furthermore, it reviews Kriging models and their designs. Finally, it reviews experiments aimed at optimization, allowing multiple random experimental outputs. This optimization may use Generalized Response Surface Methodology or Kriging combined with Mathematical Programming; the overview also covers Taguchian

Simulation-Based Optimization for Staffing and Scheduling in Call Centers

Pierre L'Ecuyer (Université de Montréal, CA)

We examine simulation-based algorithms for solving the agent scheduling problem in a multiskill call center. This problem consists in minimizing the total costs of agents under constraints on the expected service level per call type, per period, and aggregated. We discuss a solution approach that combines simulation with integer or linear programming, with cut generation. In our numerical experiments with realistic problem instances, this approach performs better than all other methods proposed previously for this problem.

We also explain why the two-step approach, which is the standard method for solving this problem, sometimes yield solutions that are highly suboptimal and inferior to those obtained by our proposed method.

Full Paper:
 http://dx.doi.org/10.1016/j.ejor.2009.01.042

Approximate Dynamic Programming in the Presence of Rare Events

Shie Mannor (Technion - Haifa, IL)

We consider approximate dynamic programming in an environment in which rare significant events occur independently of the actions selected by the controlling agent. Assuming access to a simulator, in which the rare event probabilities can be artificially altered, we introduce temporal-difference algorithms for policy evaluation, using both tabular and function approximation representations of the value function. We further introduce algorithms for policy optimization. We present empirical performance of the algorithms on a large network planning task.

Joint work of: Frank, Jordan; Mannor, Shie; Precup, Doina

Assessing Solution Quality in Stochastic Programs

David P. Morton (University of Texas - Austin, US)

Assessing whether a solution is of high quality, i.e., optimal or near optimal, is fundamental in optimization. We describe a simple Monte Carlo sampling-based procedure for assessing the quality of a candidate solution to a stochastic program. The procedure is easy to implement and widely applicable: It applies to static, and two-stage, stochastic linear programs, integer programs and convex nonlinear programs. And, it applies to multi-stage stochastic programs, given a candidate policy instead of a candidate solution. Solution quality is defined via a candidate solutionŠs optimality gap and the procedureŠs output is a confidence interval on this gap. Our simplest procedure allows significant computational improvements: (1) An asymptotically valid confidence interval can be constructed using a single-replication procedure instead of multiple replications; (2) An adaptive jackknife estimator reduces bias; (3) Common random numbers, and other variance reduction techniques, reduce sampling error; and, (4) A variant of the procedure assesses quality of a sequence of candidate solutions and can be embedded within solution algorithms requiring rigorous termination criterion.

Keywords: Stochastic programming, Monte Carlo simulation

A brief introduction to optimization via simulation

Barry L. Nelson (NW University - Evanston, US)

This presentation will provide an overview of the work that has been done by the discrete-event stochastic simulation community on optimization via simulation (which for this talk means maximizing or minimizing the expected value of some stochastic performance measure that can only be estimated via simulation). It is designed to mesh with the presentations by Kleijnen on experiment design and Chick on selection procedures, and to set up later working group discussions.

Keywords: Stochastic approximation, gradient estimation, metamodeling, random search

Modeling and optimizing under uncertainty in Vienna

Arnold Neumaier (Universität Wien, AT)

I'll report work done in Vienna on the conservative modeling of uncertainty by means of clouds, and algorithms (GRID, MCS, SNOBFIT, BBOWDA, CONVREL) that our group has developed over the years for optimizing noisy expensive functions.

Robust Planning Using Approximate Linear Programming

Marek Petrik (Univ. of Massachusetts - Amherst, US)

Developing scalable and adaptive algorithms for reasoning and acting under uncertainty is an important area in artificial intelligence. Many of these problems may be formulated as Markov decision processes and are typically solved by Approximate Dynamic Programming (ADP). While ADP has recently gained traction in many domains, the successful applications often require extensive parameter tuning in order to obtain a sufficiently small approximation error. The goal of this work is to develop ADP methods that reduce the need for extensive tuning.

In our work, we particularly focus on Approximate Linear Programming (ALP), a type of ADP. ALP has a number of theoretical advantages over other approximate dynamic programming methods, but in practice it suffers from the same performance issues as other ADP algorithms.

These issues are mostly due to a large approximation error. We analyze the approximation error and propose methods for mitigating it. First, we examine various linear program formulations and their effect on the approximation error. ALP, like other ADP methods, involves sampling, which often significantly contributes to the degradation of the solution quality. We analyze the sampling error and propose methods for minimizing it.

Joint work of: Petrik, Marek; Zilberstein, Shlomo

A Sequential Design for Approximating the Pareto Front Using an Expected Fitness Improvement Function

Thomas J. Santner (Ohio State University, US)

This talk proposes a methodology for the simultaneous optimization of multiple (expensive-to-compute) black box output functions using a surrogate modeling algorithm. The algorithm has two key features. The first feature is the use of a stochastic process-based interpolator which allows the assessment of the error of the output predictors.

The second feature is the use of a minimax fitness function as a measure of the potential improvement of adding adding training sites to the surrogate model. While the exact minimax fitness cannot be detemined, we are able to compute the expected minimax fitness given the current data. We select the next site at which each output will be computed to be that input which maximizes the expected minimax fitness. We examine two multivariate Gaussian process emulators and describe other process options. Using two measures of completeness of the solution and its spread, algorithms based on two stochastic process models, and using the expected minimax fitness and the probability of improvement (Keane, 2006) are compared in a variety of test problems.

Keywords: Surrogate model, meta model, Gaussian Process

Optimization of MRI Sampling Trajectories by Bayesian Experimental Design

Matthias Seeger (Universität des Saarlandes, DE)

We show how improved sequences for magnetic resonance imaging can be found through optimization of Bayesian design scores. Combining approximate Bayesian inference and natural image statistics with high-performance numerical computation, we propose the first Bayesian experimental design framework for this problem of high relevance to clinical practicel and brain research. Our solution requires large-scale approximate inference for dense, non-Gaussian models.

We propose a novel scalable variational inference algorithm, and show how powerful methods of numerical mathematics can be modified to compute primitives in our framework. Our approach is evaluated on raw data from a 3T MR scanner, where we achieve scan time reductions of a factor two without compromising image quality, with Cartesian and with spiral sampling trajectories.

Keywords: Bayesian Experimental Design, Variational Inference, Magnetic Resonance Imaging, Signal Processing, Compressed Sensing, Sparse Estimation, Nonlinear Design Optimization

Full Paper:
 http://people.mmci.uni-saarland.de/∼mseeger/papers/slmvar-mpi_tr.pdf

Large Scale Variational Inference and Experimental Design for Sparse Generalized Linear Models

Matthias Seeger (Universität des Saarlandes, DE)

Sparsity is a fundamental concept of modern statistics, and often the only general principle avail- able at the moment to address novel learning applications with many more variables than observations. While much progress has been made recently in the theoretical understanding and algorithmics of sparse point estimation, higher-order problems such as covariance estimation or optimal data acquisition are seldomly addressed for sparsity-favouring models, and there are virtually no algorithms for large scale applications of these. We provide novel approximate Bayesian inference algorithms for sparse generalized linear models, that can be used with hundred thousands of variables, and run orders of magnitude faster than previous algorithms in domains where either apply. By analyzing our methods and establishing some novel convexity results, we settle a long-standing open question about variational Bayesian inference for continuous variable models: the Gaussian lower bound relaxation, which has been used previously for a range of models, is proved to be a convex optimization problem, if and only if the posterior mode is found by convex programming. Our algorithms reduce to the same computational primitives than commonly used sparse estimation methods do, but require Gaussian marginal variance estimation as well. We show how the Lanczos algorithm from numerical mathematics can be employed to compute the latter.

We are interested in Bayesian experimental design here (which is mainly driven by efficient approximate inference), a powerful framework for optimizing measurement architectures of complex signals, such as natural images. Designs optimized by our Bayesian framework strongly outperform choices advocated by compressed sensing theory, and with our novel algorithms, we can scale it up to full-size images. Immediate applications of our method lie in digital photography and medical imaging.

We have applied our framework to problems of magnetic resonance imaging design and reconstruction, and part of this work appeared at a conference (Seeger et al., 2008). The present paper describes our methods in much greater generality, and most of the theory is novel. Experiments and evaluations will be given in a later paper.

Keywords: Bayesian experimental design, variational inference, sparse estimation

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/2114

Joint work of: Seeger, Matthias; Nickish, Hannes

Particle Learning and Optimization

Matt Taddy (University of Chicago, US)

In previous work, we have found that statistical modeling for the objective function provides useful guidance for sampling-based optimization schemes. In particular, full Bayesian inference – including uncertainty about model parameters – for the response surface and associated improvement statistics allows us to choose new locations for exploration based upon the posterior distribution for objective response at un-evaluated input locations. However, this type of analysis will usually require repeated runs of an MCMC sampling algorithm, as new information about the objective function is incorporated into updated model uncertainty, thus limiting the approach to applications where the cost of objective function evaluation is large relative to the cost of statistical analysis. In this talk, I will discuss a particle learning sequential Monte Carlo approach to inference for the response surface, thus allowing us to update uncertainty immediately after obtaining new data without having to re-run a longer sampling algorithm.

Regression trees will be used to model the response surface, and I will provide details for sequential updating of the associated expected improvement statistics. This framework thus allows for the robust inference of a Bayesian approach to response surface modeling, but without the computational cost of repeated MCMC.

Active Learning under Duress (ALUD)

Paul B. Kantor (Rutgers University, US)

We are interested in a group of active learning problems that are unlikely to occur in simulation but can easily occur in the real world. We call them Òactive learning under duressÓ (ALUD). These problems cannot occur with functions that are distributed in physical space, and our examples involve functions that are defined in an abstract space. The "duress" aspect of the problem arises from the fact that cases arrive very rapidly, in a time sequence determined by Mother Nature (or perhaps a wily adversary), and we must decide immediately whether to exploit or explore them.

As with all problems of this type, exploration involves a cost. And, as with oilfield examples, exploration may provide a significant positive benefit. The client organization reaps those benefits and pays those costs, and for the moment, we will pretend that it is possible to place a specific dollar value on the benefits although that is not always clear.

One kind of example involves exploration in an abstract space whose coordinates are the readings of various tests or sensors which may be applied to analyze an incoming item. In our research this item has been a cargo container arriving in port, but it could as well be a traveler arriving at a customs point, a passenger arriving at an airport, a spectator coming into a sports arena and so

forth. In this abstract model the problem is to decide whether to "thoroughly examine" the case (the equivalent of drilling for oil), to let it go, or to pass it on for further screening, together with the values of all previous test readings.

This problem can be thought of as exploration in a k-dimensional space, where k is the number of distinct sensors. Historically, it has been presented in an extrinsic form, as a branching tree of tests. The size of this tree grows super exponentially in the number of sensors. In recent work we have been able to dramatically reduce the complexity of the calculation, under the assumption that: (a) the cases of interest are extremely rare; and (b) the readings of the several tests are stochastically independent. In this formulation we have been able to successfully apply dynamic programming to solve the problem for as many as 20 sensors. In addition, making use of quite reasonable approximation schemes, we can apply it not only to sensors that produce binary readings but to sensors that produce categorical or even continuous readings.

The second application is perhaps more important, and even more difficult. This addresses the problem of monitoring millions or billions of messages in transit, and determining which may be discarded, which should be examined immediately, which should be saved for potential future analysis, and which should be submitted to further screening tests. The problem here is that, unlike the situation for the testing trees, we can not assume that we know anything about the performance characteristics of the tests that we apply. Thus we are building our scoring schemes, and learning about the incoming objects, at the same time that we are processing the data.

We conceptualize this as based in an extremely high dimensional space (D= one million or more) which is characterized, for example, by the frequency with which words or phrases occur in the messages. Although there has been application of Markov Random Fields to this process, we are not persuaded that the relationship between nearby points in this space is best understood as a random process. Our conceptual model is that every region in the space is characterized by a binomial distribution, with the probability p data items are of interest and q that they may be safely discarded. When we consider more than two possible outcomes of this becomes a multinomial distribution. We believe (if one does not believe this, then the problem seems truly hopeless) that the parameters of the distribution are smoothly varying across the space. We do not believe that the distribution is in any sense uni-modal.

This problem can in principle be characterized as a dynamic programming problem, but the size of the space is overwhelming. Therefore a variety of heuristics based on single-step (such as the knowledge gradient of Frazier and Powell), form a natural starting point. At the moment, the problem remains essentially open. We speculate that the true state space of ÒconceptsÓ which we are representing by the far-too-large space of words and phrases is of relatively low dimension (perhaps on the order of thousands). We therefore hope that joint research on reducing the dimensionality of the space, and simplifying the dynamic programming problem may be effective.

Working Group on Multiobjective Optimization

Peter Bosman, Jürgen Branke, Steve Chick, Jack Kleijnen, Joachim Kunert, Barry Nelson, Anna Syberfeldt

This working group discussed the specific challenges of sampling-based optimization in the case of more than one optimization criterion. We assumed that it is not possible to combine the different criteria into a single one, but that instead the goal is to provide the decision maker with a set of alternatives with various trade-offs. In the case of deterministic problems, one common way to address such problems is to search for a representative set of solutions that can't be improved in any objective without deteriorating at least one other objective (so-called non-dominated, Pareto-optimal, or efficient solutions). However, if a solution's objective values are stochastic, the concept of dominance is no longer obvious. If we know that the user's utility depends on aggregate information such as mean and variance, then it is possible to define dominance and optimize based on this aggregate information. The goal is then to estimate these measures efficiently. If the user's value depends on the statistical distribution of the outputs (e.g., quantiles or worst case), then this is generally not possible because in the case of multiple objectives, measures such as "quantiles" are not uniquely defined. In such cases, all the sampled values for a solution need to be considered during optimization or selection. Defining a proper metric to compare sample sets of solutions is an open research question.

The group also discussed the role of sequential simulation in multiobjective optimization. The issues identified here were: 1. Selecting the solution to be evaluated next. 2. Deciding how much simulation effort to spend on a solution. 3. Deciding when there is sufficient information so the search can be to stopped.

Working Group on Approximate Dynamic Programming

Warren Powell, Shie Mannor, Michael Fu, Xin-Rao Cao, Stephan Meisel, Marc Deisenroth, Marek Petrik

We addressed the broad challenge of designing and testing algorithms using approximate dynamic programming for different problem classes. Our discussion touched on a range of issues, including:

- How do we model a stochastic, dynamic problem? If we are going to compare algorithms for solving a problem, we have to agree on how to represent it.

But we lack a common definition for a state variable, and there are compet-
ing methods for modelling how the system evolves over time (a transition
function? One-step transition matrix? System of linear equations?)
- We identified communities which have become involved in ADP, including:
  artificial intelligence (primarily discrete state/action), engineering applica-
  tions (primarily continuous states and actions), finance, operations research.
- What are major problem classes? Some characteristics include
  - Finite/infinite horizon
  - State variables: discrete (and countable), scalar continuous, vectors (dis-
    crete and continuous)
  - Action variables: discrete (and countable), scalar continuous, vectors
    (discrete and continuous)
  - Uncertainty:
    * Deterministic
    * Stochastic
      · Expectation easy to compute
      · Expectation not computable
  - Reward structure
    * Shallow
    * Deeply nested (many moves before reward is earned)

Needless to say, the range of problem classes is quite diverse. We then dis-
cussed issues in the design and evaluation of algorithms. It is entirely possible
that people may wish to compare one aspect of an algorithm. Elements of
most ADP algorithms include:
  - What is the specification of the value function approximation?
  - How do you perform updates?
    * TD(0)? TD(lambda)?
    * Multiple inner iterations?
    * Kalman filter/recursive least squares?
    * Stochastic gradient updates?
    * ???
  - What stepsize rule do you use (if necessary)?
  - How do you handle the exploration vs. exploitation problem?

It is possible, for example, that someone may want to compare stepsize rules,
holding all other algorithmic choices constant?

There was some discussion of the specific issue of solving the exploration vs.
exploitation problem. This remains a difficult challenge in the ADP com-
munity. As of this writing, this issue is handled primarily in a heuristic
way, and primarily in the context of low-dimensional action spaces, where
"exploration" makes sense. Choosing actions to strike a balance between ex-
ploring and exploiting remains an open area of research even for optimizing
functions. Dynamic programming introduces the added dimension of biased
estimates from downstream value function approximations.

Given the design of an algorithm, there are a number of challenges that arise
in the evaluation of algorithms. Testing environments have already been
developed by the reinforcement learning and control theory communities.
The "RL Glue" environment represents an important environment which
should serve as a foundation for comparing algorithms.

Working Group on Expensive Functions

Thomas Bartz-Beielstein, Alexander Forrester, Peter Frazier, Peter Glynn, Genetha Gray, Arnold Neumaier, Günter Rudolph, Thomas Santner, Matt Taddy, Wim van Beers

This working group addressed the challenges of working with expensive, black-box computer codes. Among the issues that the group members identified as being crucial to achieving an understanding of such codes were the following.

1. Correct Identification of the Objectives of Analysis.
   Different objectives require different levels of experience with the code. Among the objectives that the working group participants had occasion to consider were
   (a) Find a (any) feasible point for code representing a complicated operating conditions.
   (b) Find the global optimum.
   (c) Find any feasible solution that is better than every previously computed function
   (d) Find entire feasible region.
   (e) Find a robust local optimum, that is to say an input with the property that "small" changes in that input does not lead to large changes in the output.
   (f) Use the code provide information about the input-output relationship for experts.
   (g) Find Pareto optimal set of inputs when there are multiple objectives
   The initial strategy one uses to successfully achieving one's goal should be multi-pronged and include the following steps.
2.1 Initial runs (sampling) of the code
   (a) Use space filling strategies (maximin Latin hypercube designs of a given run size; quasi Monte Carlo sequences)
   (b) Use fractional factorial experimental designs
   (c) Use designs that minimize integrated mean squared prediction error
2.2 Gather expert opinion
   (a) Concerning the inputs (what is the expected behavior of the output for each input when all other inputs are fixed, what are the bounds on the inputs expected in practice, what is the expected behavior of the output as a function of multiple inputs)
   (b) About the input constraints
2.3 Visualization of the input-output relationship
   (a) Nested Contour Plots
   (b) Use a 3-d "weather" map
   (c) Plot the main effect functions
   (d) Make projection pursuit plots for high dimensional inputs (XGoBI)
   (e) Use trellis or other "Deck" of contour plots

(f) Create 3-d plots with one or more sliders that allows one to vary other inputs
(g) Scatterplot matrix

Once the initial steps have been taken, one should reduce the number of inputs that are studied by screening the inputs to identify the "active" factors. There are several methods for doing this including the following.

(a) ANOVA-based sensitivity indices.
(b) Sobol' sensitivity indices
(c) Sensitivity assessment from visualizations and correlation analysis

The final tool that many group members had experience with, was the use of metamodels to supplement (or, at time, replace) expensive black-box codes. Some issues that are relevant to the use of such codes are that surrogate predictor codes are "cheap" to run compared with the black-box code. Sources of surrogate codes are:

(a) At its crudest, use "back of the envelope" models of experts
(b) Metamodels based on gradient or other information are more accurate than those that use only output information.
(c) Sometimes lower fidelity versions of the black-box code are available. In such a case, one can exploit the relationship of the output from the lower fidelity code to that of the black-box code can be exploited to improve prediction of the expensive code.
(d) Sometimes data from subsystem experiments (physical or computer) are available which can be combined with expensive code runs to improve prediction.
(e) Use model checking of the final surrogate (versus expert opinion, by cross validation) to assess its accuracy.

Working Group on Hybridization

Chun-Hung Chen, Jeff Hong, Paul Kantor, David Morton, Juta Pichitlamken, Matthias Seeger

This working group discussed possible hybridization and integration of methods from ranking and selection, optimization, and design of experiments. A summary of the discussion can be found here.

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/2117