



18th LIDS STUDENT CONFERENCE

Jan 31 - Feb 1, 2013

32-155, MIT

<http://lidsconf.mit.edu/>

Featuring

Plenaries



Corinna Cortes,
Google Research, NY



Ron Dror,
D. E. Shaw Research



Sep Kamvar,
MIT Media Lab



Michael Stonebraker,
MIT CSAIL

Student Sessions

- Learning and Optimization Algorithms
- Networks
- Optimization and Control
- Information Theory and Communication Networks
- Health Care and Biology



LIDS: Let Insights and Data Speak

Contents

Welcome to the 18 th LIDS Student Conference!!.....	3
Program at a glance	4
Invited Speakers (Titles and Abstracts)	9
Student Speakers (Titles and Abstracts)	
Learning and optimization algorithms	13
Networks	15
Optimization and Control.....	27
Information Theory and Communications	20
Health Care and Biology	22

Welcome to the 18th LIDS Student Conference (*LIDS - Let Insights and Data Speak*)!!

On behalf of the students and faculty of the Laboratory for Information and Decision Systems (LIDS) at MIT, we welcome you to the 18th LIDS Student Conference. These two days promise to be exciting and stimulating with talks spanning the diverse fields of study pursued in our lab.

This conference is an opportunity for students to present their results to their peers as well as the community-at-large. The goal is to foster communication and help spur collaboration between the various research groups at MIT.

The theme this year is "*LIDS - Let Insights and Data Speak*". Through an exciting program with student presentations, plenary speeches, and a panel discussion, we will explore how analytical approaches to systems modeling can be inspired by, and leveraged to empower, the presence of Big Data.

We are fortunate to have with us a set of distinguished plenary speakers:

Corinna Cortes (Head of Google Research, New York)

Ron Dror (Senior Research Scientist and deputy to Chief Scientist David E. Shaw, D. E. Shaw Research)

Sep Kamvar (LG Career Development Professor of Media Arts and Sciences, MIT & MIT Media Lab)

Michael Stonebraker (Adjunct Professor of Computer Science, MIT)

The conference is sponsored by the Draper Laboratory, with additional support from Prof. Peter Falb.

All talks will take place in the Stata Center (Room 32-155) at MIT. The banquet will be hosted at MIT's Faculty Club on the 6th floor of 50 Memorial Drive.

Finally, we would like to acknowledge those whose help has been invaluable in the planning and execution of this conference: Jennifer Donovan, Lynne Dell, Brian Jones, and of course our speakers and panelists. We would like to thank Draper Laboratory and Peter Falb for their sponsorship of this event. And finally, we would like to thank all of the attendees for making this event a success.

The 18th LIDS Student Conference Committee

Chairs: Kimon Drakopoulos, Ermin Wei and Kuang Xu

Noele Norris

Hamza Fawzi

Qingqing Huang

Elie Adam

Andrew Mastin

Omer Tanovic

Jianan Zhang

Yuan Shen

Henghui Lu

Schedule for Thursday, January 31, 2013

8:30-9:30 *Breakfast and Registration*

9:30-9:45 Prof. John Tsitsiklis
Opening remarks by the co-Associate Director of LIDS

9:45-10:00 Spiels of the day

10:00-11:30 ***Learning and Optimization Algorithms***
A Framework for Scalable Reward Learning from Demonstration
Buddy Michini
Algebraic Representation of Equivalence Classes of Directed Tree Models
Hajir Roozbehani
Average-Case Performance of Rollout Algorithms for Knapsack Problems
Andrew Mastin
Routing Games: Traffic Incidents and Effects of Information
Jeffrey Liu

11:30-11:45 Coffee break

11:45-12:45 ***Plenary talk***
Big Data is (at least) Four Different Problems
Prof. Michael Stonebraker

12:45-2:00 Lunch Break

Schedule for Thursday, January 31, 2013 (continued)

2:00-3:00 ***Plenary talk***
How Drugs Bind and Control Their Targets: Characterizing GPCR Signaling Using Anton, a Special-purpose Supercomputer for Molecular Dynamics Simulations
Dr. Ron Dror

3:00-3:15 Coffee break

3:15-4:45 **Networks**
Local Computation of Network Centrality
Christina Lee
Robustness of Interdependent Networks
Marzieh Parandehgheibi
Phase Transition in Random Bipartite Graphs with Given Degree Distributions
Sidhant Misra
On the Sparsity of Optimal Power Allocation for Network Localization
Wenhan Dai
Network Deconvolution: A Universal Method to Distinguish Direct Dependencies over Networks
Soheil Feizi

5:00-6:00 Panel Discussion
Let Insights and Data Speak ... to each other

Schedule for Friday, February 1, 2013

9:00-9:30 *Breakfast*

9:30-10:00 *Spiels of the day*

10:00-11:00 **Plenary talk**
Accuracy at the Top
Dr. Corinna Cortes

11:00-11:15 Coffee break

11:15-12:45 ***Optimization and Control***

A Finite Memory Encoder--Decoder Scheme for Control Under Communication Constraints

Giancarlo Baldan

Efficient Semidefinite Representations of Some Hyperbolicity Cones

James Saunderson

Random Projection Algorithms for Variational Inequalities with Set Intersection Constraints

Mengdi Wang

Minimum-violating Planning with Conflicting Specifications

Luis I. Reyes-Castro

New Lower Bounds on Nonnegative Rank Using Conic Programming

Hamza Fawzi

12:45-2:00 Lunch Break

2:00-3:00 **Plenary talk**
Organizing the World's Social Information
Sep Kamvar

Schedule for Friday, February 1, 2013 (continued)

3:00-3:15 *Coffee break*

3:15-4:45 ***Information Theory and Communication Networks***

Locally Decodable Source Coding

Ali Makhdoumi

A Proof of Threshold Saturation for Spatially-Coupled LDPC Codes on BMS Channels

Andrew Young

The Feasibility Conditions for Interference Alignment in MIMO Networks

Liangzhong Ruan

Robust Power Optimization for Wireless Network Localization

Yuan Shen

4:45-5:00 *Coffee break*

5:00-6:00 ***Health Care and Biology***

Discrimination of Biological Network Topologies Using Thermodynamic Integration

Hoda Eydgahi

The Effect of Tumor Repopulation on Optimal Fractionation Schedules in Radiation Therapy

Jagdish Ramakrishnan

Identification of Hybrid Dynamical Models of Human Motion

Ramanarayan Vasudevan

Central Sleep Apnea: Modeling, Simulation, and Identification

Ali Kazerani

6:30 Reception and Banquet (By invitation only)

Invited Speakers

Big Data is (at least) Four Different Problems

Michael Stonebraker

Adjunct Professor of Computer Science, MIT



Abstract:

Our ability to mine and analyze "Big Data" has yet to catch up to our ability to generate and collect it. Moreover "Big Data" and "Big Analytics" mean different things to different people. This talk unpacks the Big Data problem into four distinct problems, comments on trends, and compares emerging software tools in each area.

Big volumes of data with SQL analytics. The traditional data warehouse vendors support SQL analytics on very large volumes of data. In this talk, I make a few comments on where I see this market going.

Big volumes of data with complex analytics. By complex analytics, I mean data clustering, multivariate statistics, machine learning, and other complex analytics on very large amounts of data. I will explain how RDBMS vendors can be beaten by 2 orders of magnitude in this market by array DBMSs that provide real analytical scalability.

Big velocity. By this I mean being able to absorb and process a firehose of incoming data. In this market, the traditional SQL vendors are a non-starter. I will discuss alternatives including complex event processing (CEP), NoSQL and NewSQL systems.

Big diversity. Many enterprises are faced with integrating a larger and larger number of data sources with diverse data (spreadsheets, web sources, XML, traditional DBMSs). The traditional ETL products do not appear up to the challenges of this new world, and I talk about an alternate way to go.

Invited Speakers

**How Drugs Bind and Control Their Targets:
Characterizing GPCR Signaling Using Anton, a Special-
purpose Supercomputer for Molecular Dynamics
Simulations**

Ron Dror

**Senior Research Scientist at D. E. Shaw Research, and
deputy to Chief Scientist David E. Shaw**



Abstract:

Roughly one-third of all drugs act by binding to G-protein-coupled receptors (GPCRs) and either triggering or preventing receptor activation, but the process by which they do so has proven difficult to determine using either experimental or computational approaches. We recently completed a special-purpose machine, named Anton, that uses a combination of novel algorithms and application-specific hardware to accelerate molecular dynamics simulations by orders of magnitude, enabling all-atom protein simulations as long as a millisecond. Anton has made possible simulations in which drugs spontaneously associate with GPCRs to achieve bound conformations that match experimental structures almost perfectly. Simulations on Anton have also captured transitions of a GPCR between its active and inactive structures, allowing us to characterize the mechanism of receptor activation. Our results, together with complementary experimental data, suggest opportunities for the design of drugs that achieve greater specificity and control receptor signaling more precisely.

Invited Speakers

Accuracy at the Top

Corinna Cortes

Head of Google Research, NY



Abstract:

The accuracy of the items placed near the top is crucial for modern information retrieval systems, such as search engines or recommendation systems, since most users of these systems browse only the first few items. Several algorithms have been introduced in the past to optimize related notions such as precision at k , but what learning guarantees do they have?

We introduce a new notion of classification accuracy based on the top τ -quantile values of a scoring function and give a learning algorithm optimizing this criterion benefitting from margin-based guarantees. We show how training can be done by solving a set of convex optimization problems and describe a very large-scale solution based on the ADMM consensus framework. We also report the results of several experiments in the bipartite setting and compare them to those of several other algorithms seeking high precision at the top.

(Joint work with Stephen Boyd, Mehryar Mohri, and Ana Radovanovic).

Invited Speakers

Organizing the World's Social Information

Sep Kamvar

**LG Career Development Professor of Media Arts and Sciences,
MIT & MIT Media Lab**



Abstract:

In the past few years, we have seen a tremendous growth in public human communication and self-expression, through blogs, microblogs, and social networks. In addition, we are beginning to see the emergence of a social technology stack on the web, where profile and relationship information gathered by some applications can be used by other applications. In this talk I will discuss how the changing web suggests new paradigms for search and discovery, and the challenges these new paradigms pose to research in numerical linear algebra, machine learning, human-computer interaction, and programming languages.

Student Speakers

Learning and Optimization algorithms

A Framework for Scalable Reward Learning from Demonstration

Buddy Michini

Reward learning from demonstration is the task of inferring the intents or goals of an agent demonstrating a task. Inverse reinforcement learning methods utilize the Markov decision process (MDP) framework to learn rewards, but typically scale poorly since they rely on the calculation of optimal value functions. Several key modifications are made to a previously developed Bayesian nonparametric inverse reinforcement learning algorithm that avoid calculation of an optimal value function and no longer require discretization of the state or action spaces. Experimental results are given which demonstrate the ability of the resulting algorithm to scale to larger problems and learn in domains with continuous demonstrations.

Algebraic Representation of Equivalence Classes of Directed Tree Models

Hajir Roozbehani

This talk provides an algebraic representation for equivalence classes of directed tree models. It is known that a set of conditional independence (CI) relations imposed on non-singular Gaussians translates to a real algebraic variety inside the cone of positive definite covariance matrices. Thus, an equivalence class of non-singular Gaussian directed tree models is representable by a real semi-algebraic set. We provide an algebraic representation: two non-singular Gaussian directed tree models are equivalent if and only if the ideals generated by their CI (local Markov) relations are equal in the ring of polynomials on correlation coefficients. Implications on subset equivalence (for models with hidden variables) as well as equivalence of non-Gaussian tree models are further discussed.

Average-Case Performance of Rollout Algorithms for Knapsack Problems

Andrew Mastin

Rollout algorithms have demonstrated excellent performance on a variety of dynamic and discrete optimization problems. While in many cases rollout algorithms are guaranteed to perform as well as their base policies, there have been few theoretical results showing additional improvement in performance. In this paper, we perform a probabilistic analysis of the subset sum problem and knapsack problem, giving theoretical evidence that rollout algorithms perform strictly better than their base policy. Using a stochastic model from the existing literature, we analyze two rollout methods which we refer to as the consecutive rollout and exhaustive rollout, both of which employ a simple base policy. For the subset sum problem, we prove that after only a single iteration of the rollout algorithm, both methods yield at least a 30% reduction in the

expected gap between the solution value and capacity, relative to the base policy. Analogous results are shown for the knapsack problem.

Routing Games: Traffic Incidents and Effects of Information

Jeffrey Liu

The delay and congestion caused by traffic incidents costs billions of dollars in lost productivity and response costs annually. Improvements in traffic incident responses could therefore generate immense savings. A common response to traffic incidents or congestion is to communicate information to travelers through changeable message signs (CMS). Thus, quantifying the effect of CMS messages on travelers' route choices is crucial for synthesizing response strategies. In this paper, we study how access to public information about traffic conditions affects equilibrium road conditions. We assume inelastic demand, identical population preferences, and risk-neutral representative drivers. We assume that drivers choose their routes according to fully distributed strategic learning, where their information corresponds to the expected driving times and traffic incidents. We examine routing games on a parallel-link highway network and introduce an incident-dependent equilibrium concept. We provide existence conditions as well as convergence of the learning procedure under suitable conditions. We provide empirical evidence of strategic learning using publicly available traffic data in the California Bay Area.

Local Computation of Network Centrality

Christina Lee

Network centrality is a graph-based function that assigns scores to the nodes of a network. Many network centralities can be represented in terms of stationary distributions of random walks on the underlying graph -- PageRank is an example of such a network centrality. The classical power-iteration method provides a simple iterative algorithm for computing such centralities. However, each iteration of this method requires global network-wide computation which becomes extremely challenging to perform over large graphs such as World Wide Web. In this paper, we propose an efficient algorithm for approximating the network centrality of a node. It is easy to distribute and requires very little memory. Since it is a node-centric algorithm, to obtain the estimate for a single node in the network, we do not require computation over the entire network. The units of computation are simple and only require local queries of the network. We provide performance bounds on our estimates which highlight the dependence on the termination time and network properties. Our bounds yield tighter estimates for nodes with high centrality. Our analysis relies on the classical theory of positive-recurrence for countable state space Markov chain.

Robustness of Interdependent Networks

Marzieh Parandehgheibi

We consider the problem of robustness in interdependent networks. Two networks A and B are said to be interdependent if the state of network A depends on the state of network B and vice versa. We consider a cyber-physical interdependency between the networks in the sense that the nodes in network A depend on information from network B, and the nodes in network B depend on the physical output of network A. The massive blackout in Italy in 2003 was the result of such interdependency between the power grid and telecommunication network where a small failure in the power grid cascaded between the two networks and led to extensive failures in both the power grid and the telecommunication network.

We define two metrics for assessing the robustness: The first metric is the minimum number of nodes that should be removed from both networks so that all of the nodes in the network fail after the ensuing cascades, and the second one is the minimum number of nodes that should be removed from one network so that a certain given fraction of the nodes fail in the second network.

We formulate these problems, prove that it is NP-complete to evaluate these metrics optimally, and present polynomial algorithms which give suboptimal solutions.

Phase Transition in Random Bipartite Graphs with Given Degree Distributions

Sidhant Misra

The existence of a giant component in random graphs is a well studied problem. Classic results by Erdos and Renyi characterize conditions under which there exists a component of linear size in standard random graphs.

The same problem has also been studied for random graphs generated according to a given asymptotic degree sequence by Molloy and Reed. In this work we consider random multipartite graphs generated according to given degree sequences and show that there exists a phase transition phenomenon where a giant component emerges. The results are of relevance to the study of the structure of social networks, collaboration networks and the internet.

On the Sparsity of Optimal Power Allocation for Network Localization

Wenhan Dai

High-accuracy localization of mobile objects is crucial for numerous location-based applications. Power allocation among transmitting nodes not only affects network lifetime and throughput, but also determines localization accuracy. In this presentation, we formulate the power allocation problem that provides the best localization accuracy under power constraints. We first prove the sparsity property of optimal power allocation vector for network localization. Based on the sparsity, we develop optimal power allocation strategies for 2-D networks. The results provide a theoretical basis for designing transmitting node selection and power allocation strategies for network localization.

Network Deconvolution: A Universal Method to Distinguish Direct Dependencies over Networks

Soheil Feizi

Recognizing direct relationships between variables connected in a network is a pervasive problem in biological and social sciences, as correlation- and information-based networks contain numerous indirect edges due to higher-order relationships. Here, we present a universal method for inferring direct dependencies, by recognizing and removing spurious correlations due to transitive effects. We formulate the problem as the inverse of matrix transitive closure, and introduce an algorithm that solves it efficiently by exploiting eigenvector and eigenvalue decomposition and infinite-series sums. We demonstrate the effectiveness of our approach in several network applications: distinguishing direct targets in gene regulatory networks inferred from gene expression; recognizing directly-interacting amino-acid residues for protein structure prediction from sequence alignments; and distinguishing strong collaborations in co-authorship social networks using connectivity information alone. In each case, our method outperforms domain-specific methods, establishing its generality for diverse application domains in network science.

Optimization and control

A Finite Memory Encoder--Decoder Scheme for Control Under Communication Constraints

Giancarlo Baldan

The standard assumption in control theory that data can be transmitted between plant and controller without any error or delay has been replaced in recent years by more realistic models. The new area of control under communication constraints deals with designing control schemes capable of guaranteeing stability in the presence of a realistic communication channel between the plant and the controller. In this context a major role is played by the particular model chosen for the channel and the notion of stability that is required to be enforced and different types of results are available depending on the particular setting. In this research we will focus on achieving *asymptotic stability* for a discrete time one dimensional LTI plant over a *discrete memoryless* channel for which some theoretical results and practical algorithms are already available. In particular it is well known in this context that the intrinsic entropy of the plant $H = -\sum_{i: \lambda_{il} > 1} \log \lambda_{il}$ is a fundamental quantity in that stabilization is achievable iff H is lower than the capacity of the channel. In proving the sufficiency part of this statement, though, the proposed control schemes are of very little practical relevance since they are required to have infinite memory and to elaborate an unbounded number of real variables at each time. To the best of our knowledge the only exception to this situation is represented by noiseless channels for which some practically implementable algorithm achieving stability are available.

We present a new algorithm for asymptotic stabilizing a one dimensional LTI system over a discrete memoryless channel that requires only finite memory and it's of very easy implementation. The analysis of the Algorithm is non conventional and relies heavily on Birkhoff's ergodic theorem. It can be shown that the state lives on a one dimensional space where the dynamics is described by a locally expanding piecewise linear map. It is well known that this type of maps are chaotic in the sense that the long term behavior of the state can be described by an absolute continuous invariant measure and, by tuning the algorithm parameters, it is possible to adjust this invariant measure so that the system state, on average, shrinks at every time step thus ensuring asymptotic convergence.

Efficient Semidefinite Representations of Some Hyperbolicity Cones

James Saunderson

Hyperbolic programs are a family of convex optimization problems that generalize linear, second-order cone, and semidefinite programs and can be solved by interior-point methods. But it is not well understood just how general hyperbolic programs are. Indeed there are a number of (wide) open variants on the problem of whether we can reformulate any hyperbolic program (efficiently) as a semidefinite program. In this talk we mention a reasonable notion of 'reformulation' and 'efficiency' in this context and describe a natural family of hyperbolic programs that we have recently shown can be expressed efficiently in terms of semidefinite programming.

Random Projection Algorithms for Variational Inequalities with Set Intersection Constraints

Mengdi Wang

We consider the solution of strongly monotone variational inequalities of the form $F(x^*)(x - x^*) \geq 0$, for all x in X . We focus on special structures that lend themselves to sampling, such as when X is the intersection of a large number of sets, and/or F is an expected value or is the sum of a large number of component functions. We propose new methods that combine elements of incremental constraint projection and stochastic gradient. We analyze the convergence and the rate of convergence of these methods with various types of sampling schemes, and we establish a substantial rate of convergence advantage for random sampling over cyclic sampling.

Minimum-violating Planning with Conflicting Specifications

Luis I. Reyes-Castro

We consider the problem of automatic generation of control strategies for robotic vehicles given a set of high level mission specifications, such as "Vehicle x must eventually visit a target region and then return to a base", "Regions A and B must be periodically surveyed", or "None of the vehicles can enter an unsafe region". We focus on instances when all of the given specifications cannot be reached simultaneously due to their incompatibility and/or environmental constraints. We aim to find the least-violating control strategy while considering different priorities of satisfying different parts of the mission. Formally, we consider the missions given in the form of linear temporal logic formulas, each of which is assigned a reward that is earned when the formula is satisfied. Leveraging ideas from automata-based model checking, we propose an algorithm for finding an optimal control strategy that maximizes the sum of rewards earned if this control strategy is applied. We demonstrate the proposed algorithm on an illustrative case study.

Joint work with Jana Tumova, Sertac Karaman and Emilio Frazzoli.

New Lower Bounds on Nonnegative Rank Using Conic Programming

Hamza Fawzi

The nonnegative rank of an entrywise nonnegative matrix A of size $m \times n$ is the smallest integer r such that A can be written as $A = UV$ where U is $m \times r$ and V is $r \times n$ and U and V are both nonnegative. The nonnegative rank arises in different areas such as combinatorial optimization and communication complexity. Computing this quantity is NP-hard in general and it is thus important to find efficient bounding techniques especially in the context of the aforementioned applications. In this paper we propose a new lower bound on the nonnegative rank which, unlike existing lower bounds, does not explicitly rely on the matrix sparsity pattern and applies to nonnegative matrices with arbitrary support. Our lower bound is expressed as the solution of a copositive programming problem and can be relaxed to obtain polynomial-time computable lower bounds using semidefinite programming. The idea involves computing a certain nuclear

norm with nonnegativity constraints which allows to lower bound the nonnegative rank, in the same way the standard nuclear norm gives lower bounds on the standard rank. We compare our lower bound with existing ones, and we show examples of matrices where our lower bound performs better than currently known ones.

This is joint work with my advisor Prof. Pablo Parrilo.

Locally Decodable Source Coding

Ali Makhdoumi

Source coding is accomplished via the mapping of consecutive source symbols (blocks) into code blocks of fixed or variable length. In the conventional source coding in order to decode any particular source symbol, the decoder needs to read the entire encoded sequence, i.e., querying all of the encoded symbols. We study the problem of locally decodable source coding (LDSC) in which the decoder need not to read the entire encoded sequence and only a few number of queries suffices to decode any source symbol. We study both lossless and lossy source coding with a fixed number of queries as well as a growing number of queries with the source block size. We show that having a growing number of queries, one can always achieve the entropy rate for source coding. In finite length regime, we propose a lower bound on the rate of LDSC with growing number of queries. We also give a negative result on the rate of LDSC with fixed number of queries and show that the rate is one, instead of the entropy rate in conventional source coding. Finally, we demonstrate the connections to causal source coding, locally repairable codes and locally decodable codes.

A Proof of Threshold Saturation for Spatially-Coupled LDPC Codes on BMS Channels

Andrew Young

Low-density parity-check (LDPC) convolutional codes have been shown to exhibit excellent performance under low-complexity belief-propagation decoding. This phenomenon is now termed threshold saturation via spatial coupling. The underlying principle behind this appears to be very general and spatially-coupled (SC) codes have been successfully applied in numerous areas. Recently, SC regular LDPC codes have been proven to achieve capacity universally, over the class of binary memoryless symmetric (BMS) channels, under belief-propagation decoding. In Yedla et al., potential functions are used to prove that the BP threshold of SC irregular LDPC ensembles saturates, for the binary erasure channel, to the conjectured MAP threshold (known as the Maxwell threshold) of the underlying irregular ensembles. In this paper, that proof technique is generalized to BMS channels, thereby extending some results of Kudekar et al. to irregular LDPC ensembles. We also believe that this approach can be expanded to cover a wide class of graphical models whose message-passing rules are associated with a Bethe free energy.

The Feasibility Conditions for Interference Alignment in MIMO Networks

Liangzhong Ruan

Interference alignment (IA) has attracted great attention in the last few years for its breakthrough performance in interference networks. However, despite the numerous works dedicated to IA, the feasibility conditions of IA remains unclear for most network topologies. The IA feasibility analysis is challenging as the IA constraints are sets of high-degree polynomials, for which no

systematic tool to analyze the solvability conditions exists. In this work, by developing a new mathematical framework that maps the solvability of sets of polynomial equations to the linear independence of their first-order terms, we propose a sufficient condition that applies to MIMO interference networks with general configurations. We have further proved that this sufficient condition coincides with the necessary conditions under a wide range of configurations. These results further consolidate the theoretical basis of IA.

Robust Power Optimization for Wireless Network Localization

Yuan Shen

Reliable and accurate location information is essential for many emerging location-based applications in wireless networks. In network localization, power resource allocation is an important task since it affects the localization accuracy in addition to the lifetime and throughput of the network. In this work, we propose a robust power allocation formulation that guarantees the localization accuracy in the presence of parameter uncertainty. In particular, we first determine the functional properties of the localization accuracy metric and derive sequential upper and lower bounds for the worst-case localization accuracy. Built on these properties and bounds, we transform the robust formulation into a sequence of second-order cone programs (SOCPs) that yields asymptotically optimal solutions. We also develop an efficient near-optimal SOCP-based algorithm using a relaxation method. Our simulation results validate the efficiency and robustness of the proposed schemes.

Discrimination of Biological Network Topologies Using Thermodynamic Integration

Hoda Eydgahi

Signaling networks are the fundamental means by which cells transmit information used to regulate a wide range of cellular functions, such as proliferation, migration, differentiation, survival, and death. Signaling networks therefore play an important role in the development and maintenance of multicellular organisms. Due to biological and technological constraints, only minimal data is available for these large-scale biochemical networks. As a result, when comparing various model topologies, they are often deemed as indifferentiable and equiprobable due to the large degrees of freedom in each model.

We describe a Bayesian Markov Chain Monte Carlo framework that can be used for model discrimination when models produce equally good fits to data and are indistinguishable using maximum likelihood methods. We compute the Bayes factor via thermodynamic integration for two such models: the direct model and the indirect model, which are competing models of mitochondrial outer membrane permeabilization, a key step in the cell death network. We demonstrate that with 90% confidence, the direct model is 16-24 times more likely than its counterpart, indicating the greater range of parameter values that can be described by the model. The method provides a formalization of the notion of model robustness with respect to variation in parameters.

The approach proposed enables us to discriminate between competing models with different topologies. Combining this method with recent advances in single-cell measurements will enable the rigorous analysis of biochemical networks for a diverse range of cell types and diseases.

The Effect of Tumor Repopulation on Optimal Fractionation Schedules in Radiation Therapy

Jagdish Ramakrishnan

In radiation therapy, the fractionation schedule (dose delivered on each day and total number of treatment days) plays an important role in the treatment outcome. Though biological effects exist, current treatment is to little extent biologically motivated. We study the impact of different types of tumor repopulation on optimal delivery. In this work, we model tumor repopulation using popular growth models and provide an insightful analysis of the biological modeling assumptions that lead to interesting optimal fractionation schedules.

Identification of Hybrid Dynamical Models of Human Motion

Ramanarayan Vasudevan

Given the loss of freedom and the slow recovery time common to injuries arising due to falls among the elderly, the development of techniques capable of pinpointing instabilities in gait is critical. Devising an algorithm capable of detecting such deficiencies in gait requires employing

models rich enough to encapsulate the discontinuities in human motion that naturally arise due to interaction with the environment. Hybrid systems, which describe the evolution of their state both continuously via a controlled differential equation and discretely according to a control graph or a discrete input, are a class of models capable of representing such motion.

This talk begins by illustrating how a sequence of contact point enforcements along with a Lagrangian intrinsic to the human completely determines a hybrid system description for periodic human motion. The detection of contact point enforcement is then transformed into an optimal control problem for constrained switched systems. To resolve this problem, I present an approach that first relaxes the discrete switching input, performs traditional optimal control, and then projects the constructed relaxed discrete switching input back to a pure discrete switching input. The utility of this approach is illustrated by considering several examples.

Central Sleep Apnea: Modeling, Simulation, and Identification

Ali Kazerani

The respiratory chemoreflex is a major control mechanism that helps maintain blood gas homeostasis in healthy humans. Abnormalities in the chemoreflex can result in ventilatory instability, one manifestation of which defines central sleep apnea (CSA). No grey-box model yet proposed has been shown to capture the essential pathophysiology of CSA while agreeing robustly with individual patients' clinical recordings.

In this presentation, I will outline progress made toward the development of a suitable low-order grey-box model of CSA. Relevant models from the literature will first be reviewed. Our exploration of a complex, multi-system model of human physiology will then be described. Finally, a candidate low-order model will be presented, along with some early results from analysis and simulation.

The presentation concerns work in progress under the supervision of Profs. John Tsitsiklis and George Verghese.