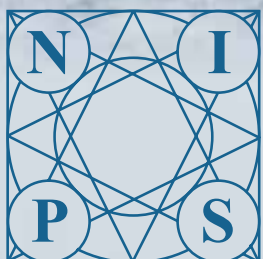
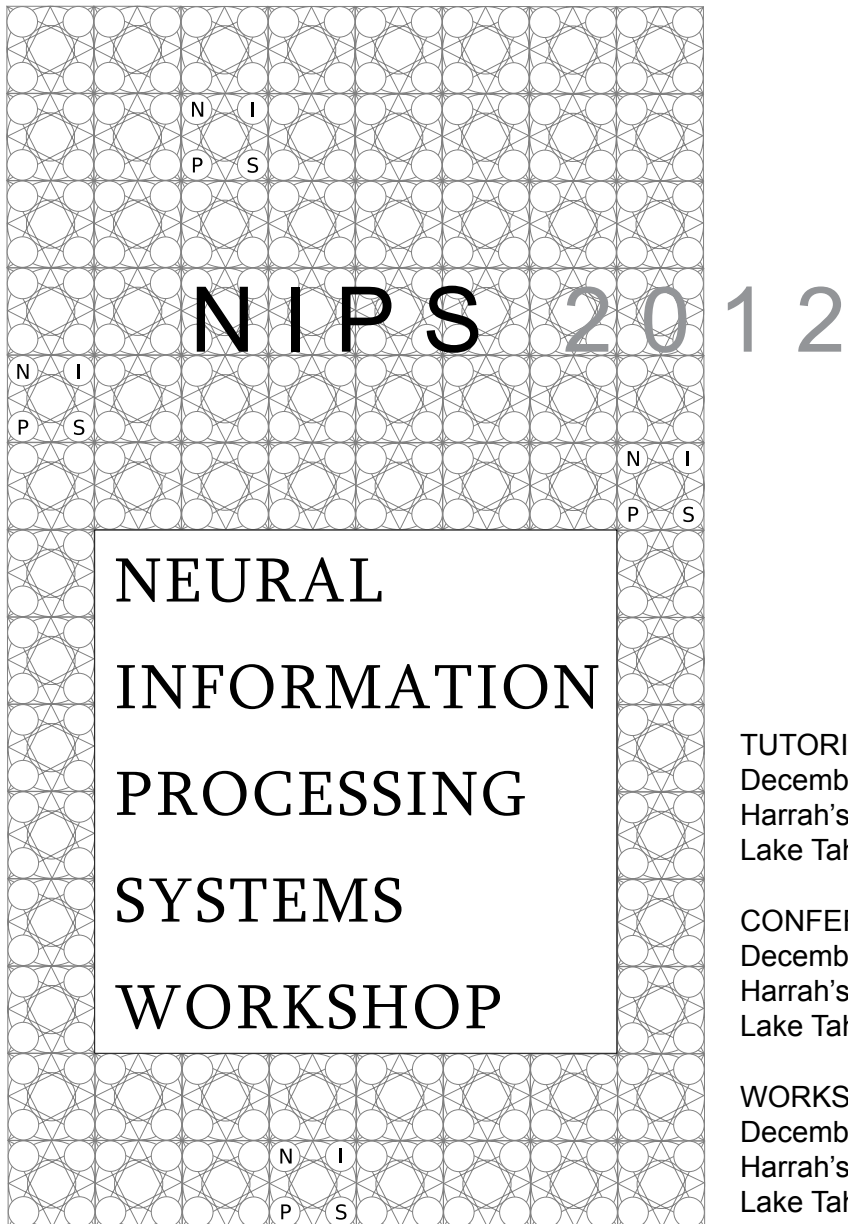


# ***2012 WORKSHOP BOOK***



Neural Information  
Processing Systems



#### TUTORIALS

December 3, 2012  
Harrah's & Harveys  
Lake Tahoe, Nevada

#### CONFERENCE SESSIONS

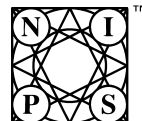
December 4 - 6, 2012  
Harrah's & Harveys  
Lake Tahoe, Nevada

#### WORKSHOPS

December 7 - 8, 2012  
Harrah's & Harveys  
Lake Tahoe, Nevada

Sponsored by the Neural Information Processing  
System Foundation, Inc

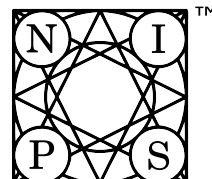
The technical program includes six invited talks and 306 accepted papers, selected from a total of 1,400 submissions considered by the program committee. Because the conference stresses interdisciplinary interactions, there are no parallel sessions. Papers presented at the conference will appear in "Advances in Neural Information Processing Systems 23," edited by Rich Zemel, John Shawe-Taylor, Peter Bartlett, Fernando Pereira and Killian Weinberger.



Neural Information  
Processing Systems  
Foundation

# TABLE OF CONTENTS

Organizing Committee	3	<b>Saturday Workshops</b>	
Program Committee	3	Algebraic Topology and Machine Learning	26
NIPS Foundation Offices and Board Members	4	Bayesian Nonparametric Models For Reliable Planning And Decision-Making Under Uncertainty	27
Core Logistics Team	4	Big Learning: Algorithms, Systems, and Tools	28
Sponsors	5	Confluence between Kernel Methods and Graphical Models	29
Maps	6	Deep Learning and Unsupervised Feature Learning	30
<b>Friday Workshops</b>		Log-Linear Models	31
Algorithmic and Statistical Approaches for Large Social Network Data Sets	9	Machine Learning Approaches to Mobile Context Awareness	32
Analysis Operator Learning vs. Dictionary Learning: Fraternal Twins in Sparse Modeling	10	MLINI - 2nd NIPS Workshop on Machine Learning and Interpretation in Neuroimaging (2-day)	33
Bayesian Optimization and Decision Making	11	Optimization for Machine Learning	34
Big Data Meets Computer Vision: First International Workshop on Large Scale Visual Recognition and Retrieval	12	Personalizing Education with Machine Learning	35
Connectomics: Opportunities and Challenges for Machine Learning	13	Perturbations, Optimization, and Statistics	36
Discrete Optimization in Machine Learning (DISCML): Structure and Scalability	14	Probabilistic Numerics	37
Human Computation for Science and Computational Sustainability	15	Probabilistic Programming: Foundations and Applications (2-day)	38
Information in Perception and Action	16	Social Choice: Theory and Practice	39
Machine Learning in Computational Biology	17	Lake Tahoe Destination and Ski Area Map	40
MLINI - 2nd NIPS Workshop on Machine Learning and Interpretation in Neuroimaging (2-day)	18		
Modern Nonparametric Methods in Machine Learning	19		
Multi-Trade-offs in Machine Learning	20		
Probabilistic Programming: Foundations and Applications (2-day)	21		
Social network and social media analysis: Methods, models and applications	22		
Spectral Algorithms for Latent Variable Models	23		
xLiTe: Cross-Lingual Technologies	24		



Neural Information  
Processing Systems  
Foundation

## ORGANIZING COMMITTEE

General Chairs: **Peter Bartlett** (Queensland University of Technology and University of California, Berkeley);  
**Fernando Pereira** (Google Research)  
Program Chairs: **Leon Bottou** (Microsoft); **Chris J.C. Burges** (Microsoft Research)  
Tutorials Chair: **Tom Griffiths** (University of California, Berkeley)  
Workshop Chairs: **Raquel Urtasun** (TTI-Chicago); **Máté Lengyel** (University of Cambridge)  
Demonstration Chair: **Thore Graepel** (Microsoft Research)  
Publications Chair and  
Electronic Proceedings Chair: **Kilian Weinberger** (Washington University in St. Louis)  
Program Manager: **Jesper Lind** (Microsoft Research)

## PROGRAM COMMITTEE

Shai Ben-David (U. Waterloo)	Noboru Murata (Waseda University)
Samy Bengio (Google)	Iain Murray (U. Edinburgh)
Matthias Bethge (Max Planck Institute Tübingen)	Sebastian Nowozin (MSR Cambridge)
Alexandre Bouchard-Côté (U. British Columbia)	Klaus Obermayer (TU Berlin)
Sebastien Bubeck (Princeton)	Aude Oliva (MIT)
Tiberio Caetano (NICTA Canberra)	Cheng Soon Ong (ETH Zürich)
Lawrence Carin (Duke)	Barak Pearlmutter (National University of Ireland Maynooth)
Ronan Collobert (IDIAP Martigny)	Doina Precup (McGill Montreal)
Marco Cuturi (Kyoto University)	Gunnar Raetsch (Memorial Sloan Kettering)
Pedro Domingos (U. Washington)	Marc'Aurelio Ranzato (Google)
Tina Eliassi-Rad (Rutgers)	Cynthia Rudin (MIT)
Rob Fergus (New York University)	Ruslan Salakhutdinov (MIT)
François Fleuret (IDIAP Martigny)	Ashutosh Saxena (Cornell)
Paolo Frasconi (U. Firenze)	Cordelia Schmid (Inria Alpes)
Amir Globerson (Hebrew University)	Fei Sha (U. Southern California)
Geoffrey Goodhill (U. of Queensland)	Shai Shalev-Shwartz (Hebrew University)
Hans-Peter Graf (NEC Labs Princeton)	Aarti Singh (CMU)
Yves Grandvalet (Université de Technologie de Compiègne)	Alex Smola (Yahoo! Research)
Kristen Grauman (U. Texas)	Masaaki Sugiyama (Tokyo Institute of Technology)
Patrick Haffner (AT&T Labs - Research)	Csaba Szepesvari (U. Alberta)
Tamir Hazan (Toyota Technological Institute Chicago)	Ambuj Tewari (U. Texas)
Katherine Heller (MIT)	Raquel Urtasun (Toyota Technological Institute Chicago)
Daniel Hsu (MSR New England)	Jean-Philippe Vert (Mines ParisTech)
Shiro Ikeda (ISM Tokyo)	Sethu Vijayakumar (Edinburgh)
Prateek Jain (MSR Bangalore)	Kilian Weinberger (U. Washington)
Thorsten Joachims (Cornell)	Lin Xiao (MSR Redmond)
Neil Lawrence (U. Sheffield)	Eric Xing (CMU)
Ping Li (Cornell)	Jieping Ye (Arizona State U.)
Phil Long (Google)	Angela Yu (UCSD)
Klaus-Robert Mueller (TU Berlin)	Kai Yu (NEC Labs Cupertino)
Remi Munos (Inria Lille)	Xiao Jin "Jerry" Zhu (U. Wisconsin)

NIPS would like to especially thank Microsoft Research for their donation of Conference Management Toolkit (CMT) software and server space.

## NIPS FOUNDATION OFFICERS & BOARD MEMBERS

President	<b>Terrence Sejnowski</b> , The Salk Institute	
Treasurer	<b>Marian Stewart Bartlett</b> , University of California, San Diego	
Secretary	<b>Michael Mozer</b> , University of Colorado, Boulder	
Legal Advisor	<b>Phil Sotel</b> , Pasadena, CA	
Executive	<b>John Lafferty</b> , Carnegie Mellon University	<b>Chris Williams</b> , University of Edinburgh
	<b>Dale Schuurmans</b> , University of Alberta, Canada	<b>Yoshua Bengio</b> , University of Montreal, Canada
	<b>Daphne Koller</b> , Stanford University	<b>Rich Zemel</b> , University of Toronto
Advisory Board	<b>Sue Becker</b> , McMaster University, Ontario, Canada	<b>Gary Blasdel</b> , Harvard Medical School
	<b>Jack Cowan</b> , University of Chicago	<b>Thomas G. Dietterich</b> , Oregon State University
	<b>Stephen Hanson</b> , Rutgers University	<b>Michael I. Jordan</b> , UC Berkeley
	<b>Michael Kearns</b> , University of Pennsylvania	<b>Scott Kirkpatrick</b> , Hebrew University, Jerusalem
	<b>Richard Lippmann</b> , MIT	<b>Todd K. Leen</b> , Oregon Graduate Institute
	<b>Bartlett Mel</b> , University of Southern California	<b>Gerald Tesauro</b> , IBM Watson Labs
	<b>Dave Touretzky</b> , Carnegie Mellon University	<b>Sebastian Thrun</b> , Stanford University
	<b>Lawrence Saul</b> , UC San Diego	<b>Sara A. Solla</b> , Northwestern University Medical School
	<b>Yair Weiss</b> , Hebrew University of Jerusalem	<b>Chris Williams</b> , University of Edinburgh
	<b>John C. Platt</b> , Microsoft Research	
	<b>John Moody</b> , International Computer Science Institute, Berkeley and Portland	
	<b>Bernhard Schölkopf</b> , Max Planck Institute for Biological Cybernetics, Tübingen	
Emeritus Members	<b>T. L. Fine</b> , Cornell University	<b>Eve Marder</b> , Brandeis University

## CORE LOGISTICS TEAM

The running of NIPS would not be possible without the help of many volunteers, students, researchers and administrators who donate their valuable time and energy to assist the conference in various ways. However, there is a core team at the Salk Institute whose tireless efforts make the conference run smoothly and efficiently every year. This year, NIPS would particularly like to acknowledge the exceptional work of:

Lee Campbell - IT Manager  
Chris Hiestand - Webmaster  
Ramona Marchand - Administrator  
Mary Ellen Perry - Executive Director

## SPONSORS

NIPS gratefully acknowledges the generosity of those individuals and organizations who have provided financial support for the NIPS 2012 conference. The financial support enabled us to sponsor student travel and participation, the outstanding student paper awards, the demonstration track and the opening buffet.



Google™



facebook

Microsoft®  
**Research**

DE Shaw & Co

TWO  SIGMA

**IBM Research**



**TOYOTA**



**YAHOO!**  
LABS

**SKYTREE**



**QUALCOMM**



# PROGRAM HIGHLIGHTS

## FRIDAY, DECEMBER 7TH

**Breakfast** - 6:30 - 8:00

See maps, Page 7

**Registration**

7:00 - 11:00 AM & 3:30 - 6:30 PM

Harvey's

### FRIDAY WORKSHOPS

All workshops run from 7:30 - 10:30 AM & 3:30 - 6:30 PM  
Coffee breaks 9:00 - 9:30 AM and 5:00 - 5:30 PM

**Algorithmic and Statistical Approaches for Large Social Network Data Sets**

Fallen Leaf & Marla Bay, Harrah's

**Analysis Operator Learning vs. Dictionary Learning: Fraternal Twins in Sparse Modeling**

Emerald Bay A, Harveys

**Bayesian Optimization and Decision Making**

Emerald Bay 1 & 2, Harveys

**Big Data Meets Computer Vision: First International Workshop on Large Scale Visual Recognition and Retrieval**

Sand Harbor 1, Harrah's

**Connectomics: Opportunities and Challenges for Machine Learning**

Emerald Bay 6, Harveys

**Discrete Optimization in Machine Learning (DISCML): Structure and Scalability**

Emerald Bay B, Harveys

**Human Computation for Science and Computational Sustainability**

Emerald Bay 4, Harveys

**Information in Perception and Action**

Emerald Bay 3, Harveys

**Machine Learning in Computational Biology**

Sand Harbor 2, Harrah's

**MLINI - 2nd NIPS Workshop on Machine Learning and Interpretation in Neuroimaging (2-day)**

Emerald Bay 5, Harveys

**Modern Nonparametric Methods in Machine Learning**

Sand Harbor 3, Harrah's

**Multi-Trade-offs in Machine Learning**

Tahoe C, Harrah's

**Probabilistic Programming: Foundations and Applications (2-day)**

Tahoe A, Harrah's

**Social Network and Social Media Analysis: Methods, Models and Applications**

Glenbrook & Emerald Bay, Harrah's

**Spectral Algorithms for Latent Variable Models**

Tahoe B, Harrah's

**xLiTe: Cross-Lingual Technologies**

Tahoe D, Harrah's

## SATURDAY, DECEMBER 8TH

**Registration**

7:00 - 11:00 AM

Harvey's 2nd Floor

### SATURDAY WOKSHOPS

All workshops run from 7:30 - 10:30 AM & 3:30 - 6:30 PM  
Coffee breaks 9:00 - 9:30 AM and 5:00 - 5:30 PM

**Algebraic Topology and Machine Learning**

Emerald Bay 6, Harveys

**Bayesian Nonparametric Models For Reliable Planning And Decision-Making Under Uncertainty**

Tahoe B, Harrah's

**Big Learning : Algorithms, Systems, and Tools**

Emerald Bay A, Harveys

**Confluence between Kernel Methods and Graphical Models**

Emerald Bay 2, Harveys

**Deep Learning and Unsupervised Feature Learning**

Emerald Bay B, Harveys

**Log-Linear Models**

Tahoe C, Harrah's

**Machine Learning Approaches to Mobile Context Awareness**

Emerald Bay 3, Harveys

**MLINI - 2nd NIPS Workshop on Machine Learning and Interpretation in Neuroimaging (2-day)**

Emerald Bay 5, Harveys

**Optimization for Machine Learning**

Fallen Leaf & Marla Bay, Harrah's

**Personalizing Education with Machine Learning**

Emerald Bay 1, Harveys

**Perturbations, Optimization, and Statistics**

Glenbrook & Emerald Bay, Harrah's

**Probabilistic Numerics**

Tahoe D, Harrah's

**Probabilistic Programming: Foundations and Applications (2-day)**

Tahoe A, Harrah's

**Social Choice: Theory and Practice**

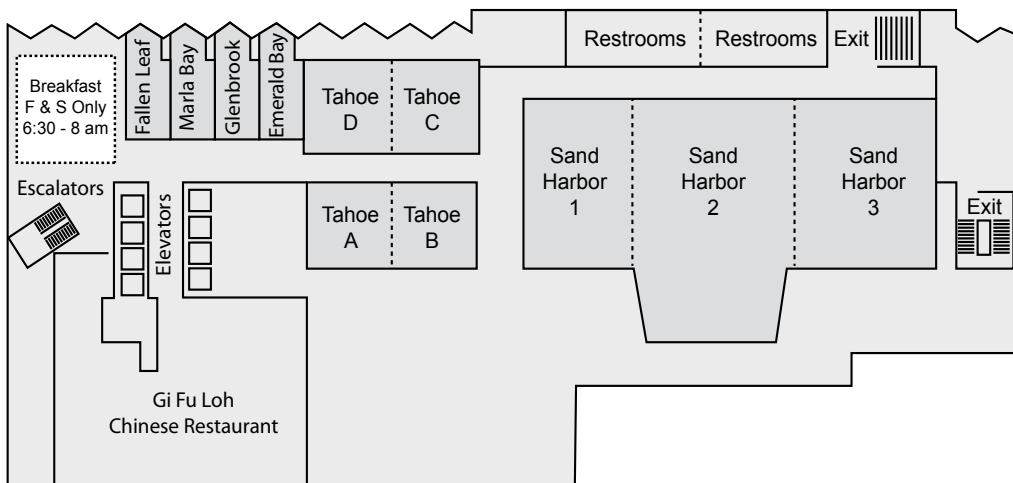
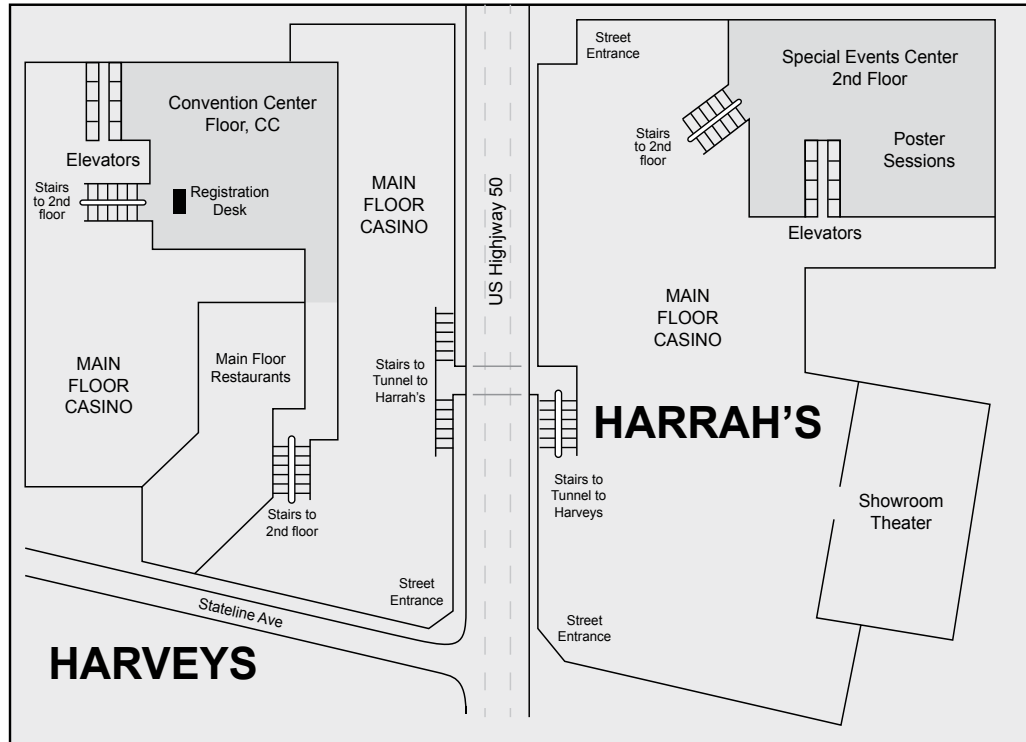
Emerald Bay 4, Harveys

**Reception**

7:00 - 11:00

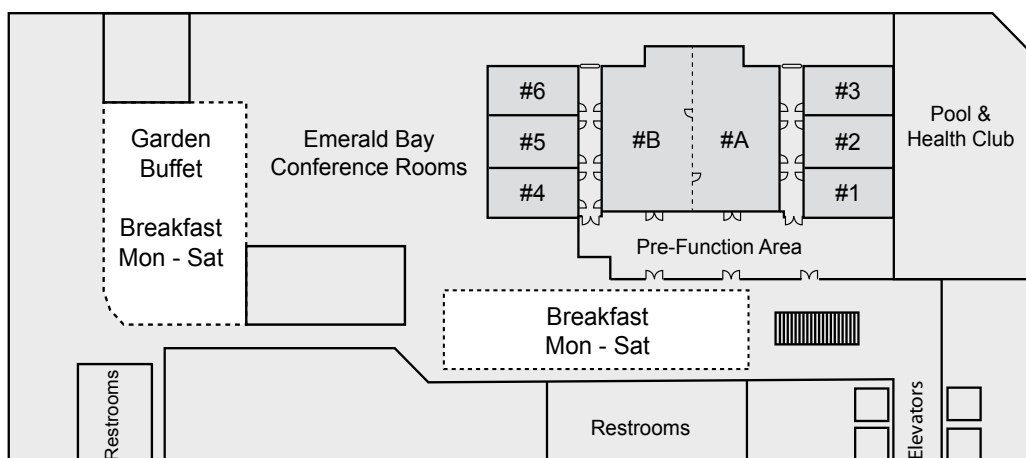
Harrah's Sand Harbor

## HARRAH'S/HARVEYS MAPS



### HARRAH'S SPECIAL EVENTS CENTER 2ND FLOOR

3 Breakfast Areas  
(dotted areas)  
1 Breakfast area at  
Top Of The Wheel  
Restaurant in  
Harvey's 12th Floor



### HARVEYS CONVENTION CENTER FLOOR, CC





# **FRIDAY WORKSHOPS**

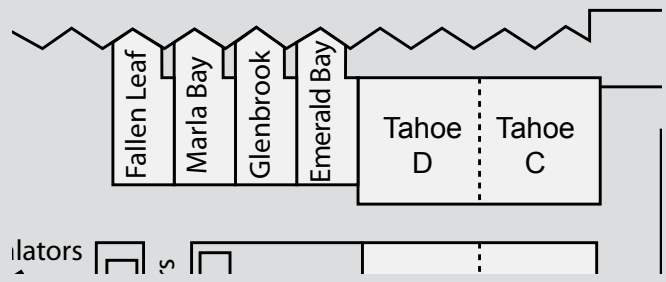
# Algorithmic and Statistical Approaches for Large Social Network Data Sets

## ORGANIZERS:

Michael Goodric	UC Irvine
Pavel Krivitsky	Penn State University
David Mount	University of Maryland
Christopher DuBois	UC Irvine
Padhraic Smyth	UC Irvine

## LOCATION:

Fallen Leaf & Marla Bay  
Harrah's Special Events Center 2nd Floor



## ABSTRACT:

Statistical models for social networks struggle with the tension between scalability - the ability to effectively and efficiently model networks with large numbers of nodes - and fidelity to social processes. Recent developments within the field have sought to address these issues in various ways, including algorithmic innovations, use of scalable latent variable models, and clever use of covariate information. This workshop will provide both a forum for presenting new innovations in this area, and a venue for debating the tradeoffs involved in differing approaches to social network modeling. The workshop will consist of a combination of invited speakers and contributed talks and posters, with ample time allowed for open discussion. Participating invited speakers include Ulrik Brandes, Carter Butts, David Eppstein, Mark Handcock, David Hunter, and David Kempe.

This workshop will be of interest to researchers working on analysis of large social network data sets, with a focus on the development of both theoretical and computational aspects of new statistical and machine learning methods for such data. Case studies and applications involving large social network data sets are also of relevance, in particular, as they impact computational and statistical issues.

Examples of specific questions of interest include:

- What are the key computational challenges involved in scaling up statistical network modeling techniques such as exponential random graph models to large networks?
- Do other techniques such as latent variable models offer useful and tractable alternatives to exponential random graph models?
- Are there new ideas from the algorithms community (in areas such as data structures and graph algorithms) that can be leveraged within network estimation algorithms?
- Can techniques from machine learning (e.g., approximate inference methods such as variational inference) be applied to statistical social network modeling?
- How can side-information (actor and edge covariates, temporal information, spatial information, textual information) be incorporated effectively into network models? How does this side-information impact computational tractability?

## WEBSITE:

<http://www.datalab.uci.edu/nips-workshop-2012/>

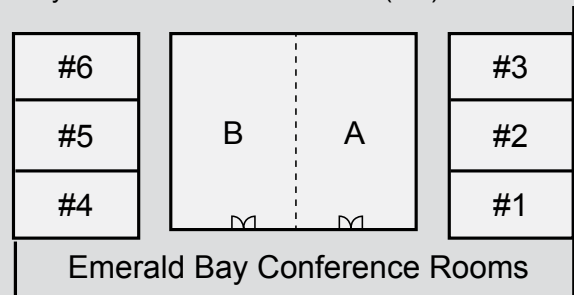
# Analysis Operator Learning vs. Dictionary Learning: Fraternal Twins in Sparse Modeling

## ORGANIZERS:

Martin Kleinsteuber	TU München, Germany
Francis Bach	INRIA
Remi Gribonval	INRIA
John Wright	Columbia University
Simon Hawe	TU München, Germany

## LOCATION:

Emerald Bay A  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Exploiting structure in data is crucial for the success of many techniques in neuroscience, machine learning, signal processing, and statistics. In this context, the fact that data of interest can be modeled via sparsity has been proven extremely valuable. As a consequence, numerous algorithms either aiming at learning sparse representations of data, or exploiting sparse representations in applications have been proposed within the machine learning and signal processing communities over the last few years. The most common way to model sparsity in data is via the so called synthesis model, also known as sparse coding. Therein, the underlying assumption is that the data can be decomposed into a linear combination of very few atoms of some dictionary. Various previous workshops and special sessions at machine learning conferences have focused on this model and its applications, as well as on algorithms for learning suitable dictionaries. In contrast to this, considerably less attention has been drawn up to now to an interesting alternative, the so called analysis model. Here, the data is mapped to a higher dimensional space by an analysis operator and the image of this mapping is assumed to be sparse. One of the most prominent examples of analysis sparsity is the total variation model in image processing. Both analysis operators and dictionaries can either be defined analytically, or learned using training samples drawn from the considered data. Learning sparse models is important since they outperform analytic ones in terms of optimal sparse representation, and allow sparse representations for classes of data where no analytical model is available. For the challenge of learning, unsupervised techniques are of major interest as they do not require labeled ground-truth data and

are independent of a specific task. There are theoretical results for the synthesis model that mathematically justify constraints on the structure of dictionaries and thus help to design learning algorithms. Nevertheless, many theoretical questions associated with learning sparse models remain, in particular for the analysis case, which is far from being fully understood. Clearly, synthesis modeling has big impact on machine learning problems like detection, classification or recognition tasks and has mainly influenced the areas of e.g. Deep Learning, or Multimodal Learning. Although the analysis model have proven advantageous over the synthesis model in regularizing inverse problems, its applicability to the aforementioned data analysis tasks has much less been investigated. The proposed workshop aims at highlighting the differences, commonalities, advantages and disadvantages of the analysis and synthesis data models. The workshop will provide a venue for discussing pros and cons of the two approaches in terms of scalability, ease of learning, and most importantly, applicability to problems in machine learning such as classification, recognition, data completion, source separation, etc. The targeted group of participants ranges from researchers in machine learning and signal processing to mathematicians. All participants of the workshop will gain a deeper understanding of the duality of the two approaches for modeling data and a clear view of which model suits best for certain applications. Moreover, further research directions will be identified that adress the issue of usability of the analysis operator approach for problems arising in Machine Learning as well as important theoretical questions related to the connection of the two fraternal twins in sparse modeling.

## WEBSITE:

<https://sites.google.com/site/dlaoplnips2012/>

**FRIDAY WORKSHOP**  
**7:30AM - 6:30PM**

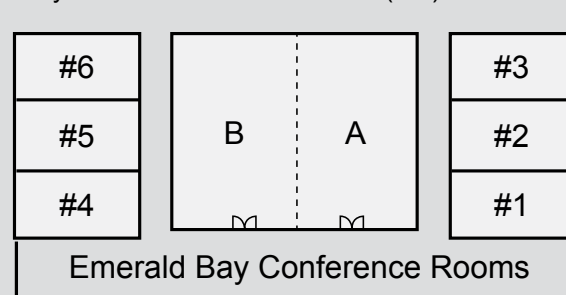
# Bayesian Optimization and Decision Making

## ORGANIZERS:

Javad Azimi	Oregon State University
Roman Garnett	Carnegie Mellon University
Frank Hutter	University of British Columbia
Shakir Mohamed	University of British Columbia

## LOCATION:

Emerald Bay 1 & 2  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Recent years have brought substantial advances in sequential decision making under uncertainty. These advances have occurred in many different communities, including several subfields of computer science, statistics, and electrical/mechanical/chemical engineering. While these communities are essentially trying to solve the same problem, they develop rather independently, using different terminology: Bayesian optimization, experimental design, bandits, active sensing, personalized recommender systems, automatic algorithm configuration, reinforcement learning, and so on. Some communities focus more on theoretical aspects while others' expertise is on real-world applications. This workshop aims to bring researchers from these communities together to facilitate cross-fertilization by discussing challenges, findings, and sharing data. This

workshop follows last year's NIPS workshop "Bayesian optimization, experimental design and bandits: Theory and applications", one of the most-attended workshops in 2011. This year we plan to focus somewhat more on real-world applications, to bridge the gap between theory and practice. Specifically, we plan to have a panel discussion on real-world and industrial applications of Bayesian optimization and an increased focus on real-world applications in the invited talks (covering hyperparameter tuning, configuration of algorithms for solving hard combinatorial problems, energy optimization, and optimization of MCMC). Similar to last year, we expect to highlight the most beneficial research directions and unify the whole community by setting up this workshop.

## WEBSITE:

<http://javad-azimi.com/nips2012ws/>

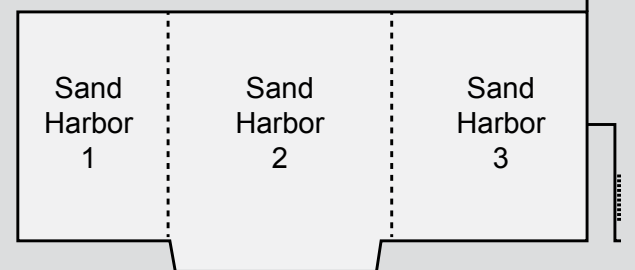
# Big Data Meets Computer Vision: First International Workshop on Large Scale Visual Recognition and Retrieval

## ORGANIZERS:

Jia Deng	Stanford University
Samy Bengio	Google
Yuanqing Lin	NEC Labs America
Fei Fei Li	Stanford University

## LOCATION:

Sand Harbor 1  
Harrah's Special Events Center 2nd Floor



## ABSTRACT:

The emergence of "big data" has brought about a paradigm shift throughout computer science. Computer vision is no exception. The explosion of images and videos on the Internet and the availability of large amounts of annotated data have created unprecedented opportunities and fundamental challenges on scaling up computer vision.

Over the past few years, machine learning on big data has become a thriving field with a plethora of theories and tools developed. Meanwhile, large scale vision has also attracted increasing attention in the computer vision community. This workshop aims to bring closer researchers in large scale machine learning and large scale vision to foster cross-talk between the two fields. The goal is to encourage machine learning researchers to work on large scale vision problems, to inform computer vision researchers about new developments on large scale learning, and to identify unique challenges and opportunities.

This workshop will focus on two distinct yet closely related vision problems: recognition and retrieval. Both are inherently large scale. In particular, both must handle high dimensional features (hundreds of thousands to millions), a large variety of visual classes (tens of thousands to millions), and a large number of examples (millions to billions).

This workshop will consist of invited talks, panels, discussions, and paper submissions including, but not limited to, the following topics:

- State of the field: What really defines large scale vision? How does it differ from traditional vision research? What are its unique challenges for large scale learning?
- Indexing algorithms and data structures: How do we efficiently

find similar features/images/classes from a large collection, a key operation in both recognition and retrieval?

- Semi-supervised/unsupervised learning: Large scale data comes with different levels of supervision, ranging from fully labeled and quality controlled to completely unlabeled. How do we make use of such data?
- Metric learning: Retrieval visually similar images/objects requires learning a similarity metric. How do we learn a good metric from a large amount of data?
- Visual models and feature representations: What is a good feature representation? How do we model and represent images/videos to handle tens of thousands of fine-grained visual classes?
- Exploiting semantic structures: How do we exploit the rich semantic relations between visual categories to handle a large number of classes?
- Transfer learning: How do we handle new visual classes (objects/scenes/activities) after having learned a large number of them? How do we transfer knowledge using the semantic relations between classes?
- Optimization techniques: How do we perform learning with training data that do not fit into memory? How do we parallelize learning?
- Datasets issues: What is a good large scale dataset? How should we construct datasets? How do we avoid dataset bias?
- Systems and infrastructure: How do we design and develop libraries and tools to facilitate large scale vision research? What infrastructure do we need?

The target audience of this workshop includes industry and academic researchers interested in machine learning, computer vision, multimedia, and related fields.

## WEBSITE:

<https://sites.google.com/site/bigvision2012/>

**FRIDAY WORKSHOP**  
**7:30AM - 6:30PM**

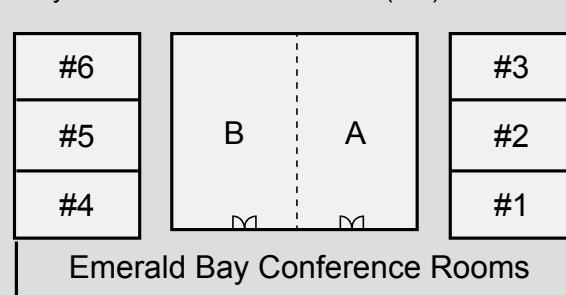
# Connectomics: Opportunities and Challenges for Machine Learning

## ORGANIZERS:

Viren Jain                      Howard Hughes Medical Institute  
Moritz Helmstaedter        Max Planck Institute of Neurobiology

## LOCATION:

Emerald Bay 6  
Harveys Convention Center Floor (CC)



## ABSTRACT:

The “wiring diagram” of essentially all nervous systems remains unknown due to the extreme difficulty of measuring detailed patterns of synaptic connectivity of entire neural circuits. At this point, the major bottleneck is in the analysis of tera or peta-voxel 3d electron microscopy image data in which neuronal processes need to be traced and synapses localized in order for connectivity information to be inferred. This presents an opportunity for machine learning and machine perception to have a fundamental impact on advances in neurobiology. However, it also presents a major challenge, as existing machine learning methods fall short of solving the problem.

The goal of this workshop is to bring together researchers in machine learning and neuroscience to discuss progress and remaining challenges in this exciting and rapidly evolving field. We aim to attract machine learning and computer vision specialists interested in learning about a new problem, as well as computational neuroscientists at NIPS who may be interested in modeling connectivity data. We will discuss the release of public datasets and competitions that may facilitate further activity in this area. We expect the workshop to result in a significant increase in the scope of ideas and people engaged in this field. As NIPS bridges both the neuroscience and computer science communities, it is the ideal venue for this type of workshop.

## WEBSITE:

<http://www.nips2012connectomics.net>



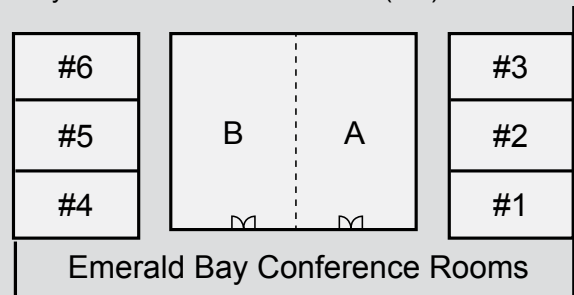
# Discrete Optimization in Machine Learning (DISCML): Structure and Scalability

## ORGANIZERS:

Stefanie Jegelka	UC Berkeley
Andreas Krause	ETH Zurich
Jeff Bilmes	University of Washington
Pradeep Ravikumar	University of Texas, Austin

## LOCATION:

Emerald Bay B  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Optimization problems with discrete solutions (e.g., combinatorial optimization) are becoming increasingly important in machine learning. The core of statistical machine learning is to infer conclusions from data, and when the variables underlying the data are discrete, both the tasks of inferring the model from data, as well as performing predictions using the estimated model are discrete optimization problems. Two factors complicate matters: first, many discrete problems are in general computationally hard, and second, machine learning applications often demand solving such problems at very large scales.

The focus of this year's workshop lies on structures that enable scalability. Examples of important structures include sparse graphs, the marginal polytope, and submodularity. Which properties of the problem make it possible to still efficiently obtain exact or decent approximate solutions? What are the challenges posed by parallel and distributed processing? Which discrete problems in machine learning are in need of more scalable algorithms? How can we make discrete algorithms scalable while retaining quality? Some heuristics perform well but as of yet are devoid of a theoretical foundation; what explains such good behavior?

## WEBSITE:

<http://discml.cc>

**FRIDAY WORKSHOP**  
**7:30AM - 6:30PM**

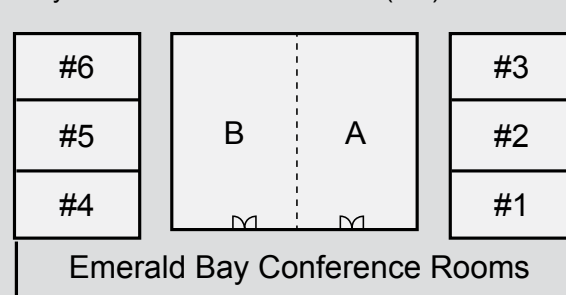
# Human Computation for Science and Computational Sustainability

## ORGANIZERS:

Theo Damoulas	Cornell University
Thomas Dietterich	Oregon State University
Edith Law	Carnegie Mellon University
Serge Belongie	UC San Diego

## LOCATION:

Emerald Bay 4  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Researchers in several scientific and sustainability fields have recently achieved exciting results by involving the general public in the acquisition of scientific data and the solution of challenging computational problems. One example is the eBird project ([www.ebird.org](http://www.ebird.org)) of the Cornell Lab of Ornithology, where field observations uploaded by bird enthusiasts are providing continent-scale data on bird distributions that support the development and testing of hypotheses about bird migration. Another example is the FoldIt project ([www.fold.it](http://www.fold.it)), where volunteers interacting with the FoldIt software have been able to solve the 3D structures of several biologically important proteins.

Despite these early successes, the involvement of the general public in these efforts poses many challenges for machine learning. Human observers can vary hugely in their degree of expertise. They conduct observations when and where they see fit, rather than following carefully designed experimental protocols. Paid participants (e.g., from Amazon Mechanical Turk) may not follow the rules or may even deliberately mislead the investigators.

A related challenge is that problem instances presented to human participants can vary in difficulty. Some instances (e.g., of visual tasks) may be impossible for most people to solve. This leads to a bias toward easy instances, which can confuse learning algorithms.

A third issue with crowdsourcing is that in many of these problems, there is no available ground truth because the true quantities of interest are only indirectly observed. For example, the BirdCast project seeks to model the migration of birds. However, the eBird reports only provide observations of birds on or near the ground, rather than migratory flight (which occurs predominantly at night). In such situations, it is hard to evaluate the accuracy of the learned models, because predictive accuracy does not guarantee that the values of latent variables are correct or that the model is identifiable.

This workshop will bring together researchers at the interface of machine learning, citizen science, and human computation. The goals of the workshop are i) to identify common problems, ii) to propose benchmark datasets, common practices and improved methodologies for dealing with

such phenomena, iii) to identify methods for evaluating such models in the absence of ground truth, iv) to share approaches for implementing and deploying citizen science and human computation projects in scientific and sustainability domains, and v) to foster new connections between the scientific, sustainability, and human computation research communities.

There will be two awards (250\$ book vouchers) for Best Contribution for the oral and/or poster presentations sponsored by the Institute for Computational Sustainability ([www.cis.cornell.edu/ics](http://www.cis.cornell.edu/ics))

We welcome submissions\* related to (but not limited to) the following topics:

- Biases, probabilistic observation processes, noise processes, and other imperfections in citizen science and human computation
- Novel citizen science and human computation projects in science and sustainability
- Human-machine interactions and human computation in science and sustainability
- Novel modeling paradigms for citizen science and human computation projects in science and sustainability
- Methods for recruitment, retention, and modeling of human participants
- Dataset shift and domain adaptation methods for citizen science and human computation projects
- Spatio-temporal active learning and general inference techniques on citizen science and human computation projects

\*Every submission should include both a paper and a poster. The paper should be in standard nips format (final format, not a blind-author submission), up to 4 pages maximum and the poster in portrait format and up to 33" x 47" (A0) dimensions.

Expected outcomes of this workshop include: a list of open problems, acquaintance with relevant work between scientific domains, a 5-year research roadmap, emerging collaborations between participants, and attracting more people to work in Computational Sustainability and Human Computation. In addition, we will explore avenues for organizing a future machine learning competition, on a large-scale HC problem, to foster and strengthen the community around a prototypical HC effort.

## WEBSITE:

<http://www.cs.cornell.edu/~damoulas/Site/HCSCS.html>

**FRIDAY WORKSHOP**  
**7:30AM - 6:30PM**

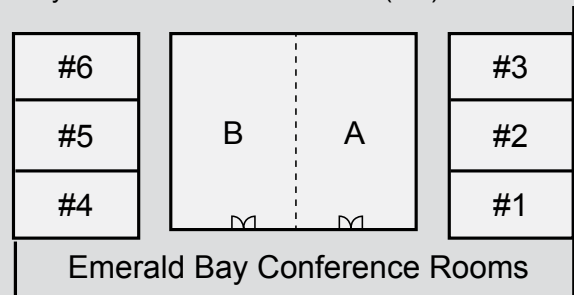
# Information in Perception and Action

## ORGANIZERS:

Naftali Tishby	Hebrew University Jerusalem
Daniel Polani	University of Hertfordshire
Tobias Jung	University of Liege

## LOCATION:

Emerald Bay 3  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Since its inception for describing the laws of communication in the 1940's, information theory has been considered in fields beyond its original application area and, in particular, it was long attempted to utilize it for the description of intelligent agents. Already Attneave (1954) and Barlow (1961) suspected that neural information processing might follow principles of information theory and Laughlin (1998) demonstrated that information processing comes at a high metabolic cost; this implies that there would be evolutionary pressure pushing organismic information processing towards the optimal levels of data throughput predicted by information theory. This becomes particularly interesting

when one considers the whole perception-action cycle, including feedback. In the last decade, significant progress has been made in this direction, linking information theory and control. The ensuing insights allow to address a large range of fundamental questions pertaining not only to the perception-action cycle, but to general issues of intelligence, and allow to solve classical problems of AI and machine learning in a novel way.

The workshop will present recent work on progress in AI, machine learning, control, as well as biologically plausible cognitive modeling, that is based on information theory.

## WEBSITE:

<http://www.montefiore.ulg.ac.be/~tjung/nips12workshop>

**FRIDAY WORKSHOP**  
**7:30AM - 6:30PM**

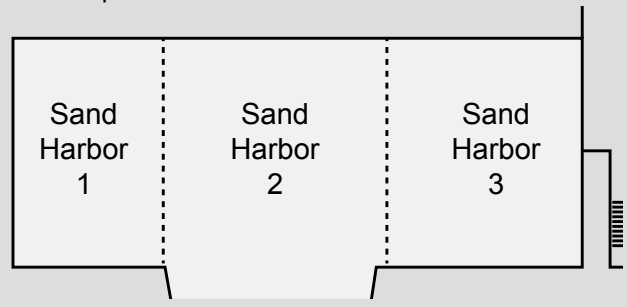
# Machine Learning in Computational Biology

## ORGANIZERS:

Jean-Philippe Vert	ParisTech
Anna Goldenberg	University of Toronto
Christina Leslie	Memorial Sloan-Kettering Cancer Center

## LOCATION:

Sand Harbor 2  
Harrah's Special Events Center 2nd Floor



## ABSTRACT:

The field of computational biology has seen dramatic growth over the past few years, both in terms of new available data, new scientific questions, and new challenges for learning and inference. In particular, biological data are often relationally structured and highly diverse, well-suited to approaches that combine multiple weak evidence from heterogeneous sources. These data may include sequenced genomes of a variety of organisms, gene expression data from multiple technologies, protein expression data, protein sequence and 3D structural data, protein interactions, gene ontology and pathway databases, genetic variation data (such as SNPs), and an enormous amount of textual data in the biological and medical literature. New types of scientific and clinical problems require the development of novel supervised and unsupervised learning methods that can use these growing resources. Furthermore, next generation sequencing technologies are yielding terabyte scale data sets that require novel algorithmic solutions.

The goal of this workshop is to present emerging problems and machine learning techniques in computational biology. We will invite several speakers from the biology/bioinformatics community who will present current research problems in bioinformatics, and we will invite contributed talks on novel learning approaches in computational biology. We encourage contributions describing either progress on new bioinformatics problems or work on established problems using methods that are substantially different from standard approaches. Kernel methods, graphical models, feature selection, and other techniques applied to relevant bioinformatics problems would all be appropriate for the workshop. The targeted audience are people with interest in learning and applications to relevant problems from the life sciences.

## WEBSITE:

<http://www.mlcg.org>

# MLINI - 2nd NIPS Workshop on Machine Learning and Interpretation in Neuroimaging (2-day)

## ORGANIZERS:

Georg Langs	Medical University of Vienna
Irina Rish	IBM Research
Guillermo Cecchi	IBM Research
Brian Murphy	Carnegie Mellon University
Bjoern Menze	ETH Zurich
Kai-min Chang	Carnegie Mellon University
Moritz Grosse-Wentrup	Max Planck Institute

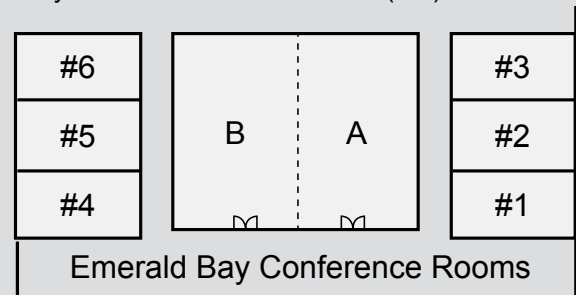
## ABSTRACT:

A workshop on the topic of machine learning approaches in neuroscience and neuroimaging. We believe that both machine learning and neuroimaging can learn from each other as the two communities overlap and enter an intense exchange of ideas and research questions. Methodological developments in machine learning spurn novel paradigms in neuroimaging, neuroscience motivates methodological advances in computational analysis. In this context many controversies and open questions exist. The goal of the workshop is to pinpoint these issues, sketch future directions, and tackle open questions in the light of novel methodology.

The first workshop of this series at NIPS 2011 built upon earlier events in 2006 and 2008. Last year's workshop included many invited speakers, and was centered around two panel discussions, during which 2 questions were discussed: the interpretability of machine learning findings, and the shift of paradigms in the neuroscience community. The discussion was inspiring, and made clear, that there is a tremendous amount the two communities can learn from each other benefiting from communication across the disciplines.

## LOCATION:

Emerald Bay 5  
Harveys Convention Center Floor (CC)



The aim of the workshop is to offer a forum for the overlap of these communities. Besides interpretation, and the shift of paradigms, many open questions remain. Among them:

- How suitable are MVPA and inference methods for brain mapping?
- How can we assess the specificity and sensitivity?
- What is the role of decoding vs. embedded or separate feature selection?
- How can we use these approaches for a flexible and useful representation of neuroimaging data?
- What can we accomplish with generative vs. discriminative modelling?
- Can and should the Machine Learning community provide a standard repertoire of methods for the Neuroimaging community to use (e.g. in choosing a classifier)?

## WEBSITE:

<https://sites.google.com/site/nipsmlini2012/>

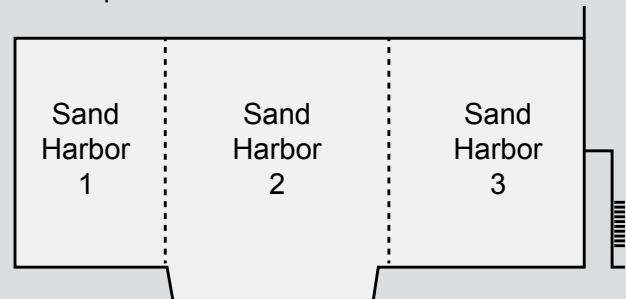
# Modern Nonparametric Methods in Machine Learning

## ORGANIZERS:

Sivaraman Balakrishnan	Carnegie Mellon University
Arthur Gretton	Gatsby Unit, UCL
Mladen Kolar	Carnegie Mellon University
John Lafferty	University of Chicago
Han Liu	Princeton University
Tong Zhang	Rutgers University

## LOCATION:

Sand Harbor 3  
Harrah's Special Events Center 2nd Floor



## ABSTRACT:

The objective of this workshop is to bring together practitioners and theoreticians who are interested in developing scalable and principled nonparametric learning algorithms for analyzing complex and large-scale datasets. The workshop will communicate the newest research results and attack several important bottlenecks of nonparametric learning by exploring (i) new models and methods that enable high-dimensional nonparametric learning, (ii) new computational techniques that enable scalable nonparametric learning in online and parallel fashion, and (iii) new statistical theory that characterizes the performance and information-theoretic limits of nonparametric learning algorithms. The expected goals of this workshop include (i) reporting the state-of-the-art of modern nonparametrics, (ii) identifying major challenges and setting up the frontiers for nonparametric methods, (iii) connecting different disjoint communities in machine learning and statistics. The targeted application areas include genomics, cognitive neuroscience, climate science, astrophysics, and natural language processing.

Modern data acquisition routinely produces massive and complex datasets, including chip data from high throughput genomic experiments, image data from functional Magnetic Resonance Imaging (fMRI), proteomic data from tandem mass spectrometry analysis, and climate data from geographically distributed data centers. Existing high dimensional theories and learning algorithms rely heavily on parametric models, which assume the data come from an underlying distribution (e.g. Gaussian or linear models) that can be characterized by a finite number of parameters. If these assumptions are correct, accurate and precise estimates can be expected. However, given the increasing complexity of modern scientific datasets, conclusions inferred under these restrictive assumptions can be misleading. To handle this challenge, this workshop focuses on nonparametric methods, which directly conduct inference in infinite-dimensional spaces and thus are powerful enough to capture the subtleties in most modern applications.

We are targeting submissions in a variety of areas. Potential topics include, but are not limited to, the following areas where high dimensional nonparametric methods have found past success:

1. Nonparametric graphical models are a flexible way to model continuous distributions. For example, copulas can be used to separate the dependency structure between random variables from their marginal distributions (Liu et al. 2009). Fully nonparametric model of networks can be obtained using kernel density estimation and restricting the graphs to trees and forests (Liu et al. 2011).
2. Causal inference using kernel-based conditional independence testing is a nonparametric method, which improves a lot over previous approaches to estimate or test for conditional independence (Zhang et al. 2012).
3. Sparse additive models are used in many applications where linear regression models do not provide enough flexibility (Lin and Zhang, 2006), (Koltchinskii and Yuan, 2010), (Huang et al. 2010), (Ravikumar et al. 2009), (Meier et al. 2009).
4. Nonparametric methods are used to consistently estimate a large class of divergence measures, which have a wide range of applications (Póczos and Schneider, 2011).
5. Recently sparse matrix decompositions (Witten et al., 2009) were proposed as exploratory data analysis tools for high dimensional genomic data. Motivated by the need for additional modelling flexibility, sparse nonparametric generalizations of these matrix decompositions have been introduced (Balakrishnan et al., 2012).
6. Nonparametric learning promises flexibility, where flexible methods minimize assumptions such as linearity and Gaussianity that are often made only for convenience, or lack of alternatives. However, nonparametric estimation often comes with increased computational demands. To develop algorithms that are applicable on large-scale data, we need to take advantage of parallel computation. Promising parallel computing techniques include GPU programming, multi-core computing, and cloud computing.

## WEBSITE:

<http://sites.google.com/site/nips2012modernnonparametric>



# Multi-Trade-offs in Machine Learning

## ORGANIZERS:

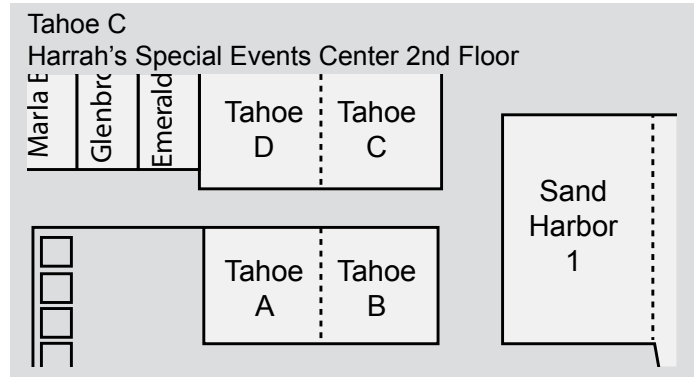
Yevgeny Seldin	Max Planck Institute
Guy Lever	University College London
John Shawe-Taylor	University College London
Nicolò Cesa-Bianchi	Università degli Studi di Milano
Koby Crammer	The Technion
Francois Laviolette	Université Laval
Gabor Lugosi	Pompeu Fabra University
Peter Bartlett	UC Berkeley

## ABSTRACT:

One of the main practical goals of machine learning is to identify relevant trade-offs in different problems, formalize, and solve them. We have already achieved fairly good progress in addressing individual trade-offs, such as model order selection or exploration-exploitation. In this workshop we would like to focus on problems that involve more than one trade-off simultaneously. We are interested both in practical problems where “multi-trade-offs” arise and in theoretical approaches to their solution. Obviously, many problems in life cannot be reduced to a single trade-off and it is highly important to improve our ability to address multiple trade-offs simultaneously. Below we provide several examples of situations, where multiple trade-offs arise simultaneously. The goal of the examples is to provide a starting point for a discussion, but they are not limiting the scope and any other multi-trade-off problem is welcome to be discussed at the workshop.

Multi-trade-offs arise naturally in interaction between multiple learning systems or when a learning system faces multiple tasks simultaneously; especially when the systems or tasks share common resources, such as CPU time, memory, sensors, robot body, and so on. For a concrete example, imagine a robot riding a bicycle and balancing a pole. Each task individually (cycling and pole balancing) can be modeled as a separate optimization problem, but their solutions have to be coordinated, since they share robot resources and robot body. More generally, each learning system or system

## LOCATION:



component has its own internal trade-offs, which have to be balanced against trade-offs of other systems, whereas shared resources introduce external trade-offs that enforce cooperation. The complexity of interaction can vary from independent systems sharing common resources to systems with various degrees of relation between their inputs and tasks. In multi-agent systems communication between the agents introduces an additional trade-off.

We are also interested in multi-trade-offs that arise within individual systems. For example, model order selection and computational complexity, or model order selection and exploration-exploitation. For a specific example of this type of problems, imagine a system for real-time prediction of the location of a ball in table tennis. This system has to balance between at least three objectives that interact in a non-trivial manner: (1) complexity of the model of flight trajectory, (2) statistical reliability of the model, (3) computational requirements. Complex models can potentially provide better predictions, but can also lead to overfitting (trade-off between (1) and (2)) and are computationally more demanding. At the same time, there is also a trade-off between having fast crude predictions or slower, but more precise estimations (trade-off between (3) and (1)+(2)).

Despite the complex nature of multi-trade-offs, there is still hope that they can be formulated as convex problems, at least in some situations.

## WEBSITE:

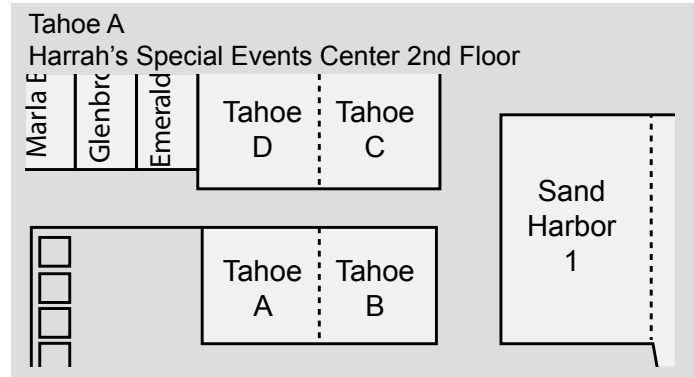
<https://sites.google.com/site/multitradeoffs2012/>

# Probabilistic Programming: Foundations and Applications (2-day)

## ORGANIZERS:

Vikash Mansinghka    MIT and Prior Knowledge, Inc  
Dan Roy                University of Cambridge  
Noah Goodman        Stanford University

## LOCATION:



## ABSTRACT:

An intensive, two-day workshop on PROBABILISTIC PROGRAMMING, with contributed and invited talks, poster sessions, demos, and discussions.

Probabilistic models and inference algorithms have become standard tools for interpreting ambiguous, noisy data and building systems that learn from their experience. However, even simple probabilistic models can require significant effort and specialized expertise to develop and use, frequently involving custom mathematics, algorithm design and software development. State-of-the-art models from Bayesian statistics, artificial intelligence and cognitive science --- especially those involving distributions over infinite data structures, relational structures, worlds with unknown numbers of objects, rich causal simulations of physics and psychology, and the reasoning processes of other agents - can be difficult to even specify formally, let alone in a machine-executable fashion.

PROBABILISTIC PROGRAMMING aims to close this gap, making variations on commonly-used probabilistic models far easier to develop and use, and pointing the way towards entirely new types of models and inference. The central idea is to represent probabilistic models using ideas from programming, including functional, imperative, and logic-based languages. Most probabilistic programming systems represent distributions algorithmically, in terms of a programming language plus primitives for stochastic choice; some even support inference over Turing-universal languages. Compared with representations of models in terms of their graphical-model structure, these representation languages are often significantly more flexible, but still support the development of general-purpose inference algorithms.

The workshop will cover, and welcomes submissions about, all aspects of probabilistic programming. Some questions of particular interest include:

1. What real-world problems can be solved with probabilistic programming systems today? How much problem-specific customization/optimization is needed? Where is general-purpose inference effective?
2. What does the probabilistic programming perspective, and in particular the representation of probabilistic models and inference procedures as algorithmic processes, reveal about the computability and complexity of Bayesian inference? When can theory guide the design and use of probabilistic programming systems?
3. How can we teach people to write probabilistic programs that work well, without having to teach them how to build an inference engine first? What programming styles support tractability of inference?
4. How can central ideas from software engineering --- including debuggers, validation tools, style checkers, program analyses, reusable libraries, and profilers --- help probabilistic programmers and modelers? Which of these tools can be built for probabilistic programs, or help us build probabilistic programming systems?
5. What new directions in AI, statistics, and cognitive science would be enabled if we could handle models that took hundreds or thousands of lines of probabilistic code to write?

Confirmed Keynote Speakers: - Josh Tenenbaum (MIT), Stuart Russell (UC Berkeley), Christopher Bishop (Microsoft Research)

## WEBSITE:

[http://probabilistic-programming.org/wiki/NIPS\\*2012\\_Workshop](http://probabilistic-programming.org/wiki/NIPS*2012_Workshop)

FRIDAY WORKSHOP  
7:30AM - 6:30PM

# Social network and social media analysis: Methods, models and applications

## ORGANIZERS:

Georg Langs	Medical University of Vienna
Irina Rish	IBM Research
Guillermo Cecchi	IBM Research
Brian Murphy	Carnegie Mellon University
Bjoern Menze	ETH Zurich
Kai-min Chang	Carnegie Mellon University
Moritz Grosse-Wentrup	Max Planck Institute

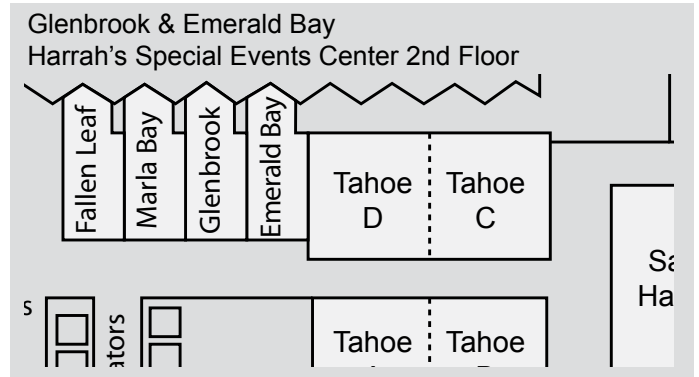
## ABSTRACT:

Modern technology, including the World Wide Web, sensor networks, and high-throughput genetic sequencing, has completely transformed the scale and concept of data in the sciences. Data collections for a number of systems of interest have grown large and heterogeneous, and a crucial subset of the data is often represented as a collection of graphs together with node and edge attributes. Thus, the analysis and modeling of large, complex, real-world networks has become necessary in the study of phenomena across the diverse set of social, technological, and natural worlds. The aim of this workshop is to bring together researchers with these diverse sets of backgrounds and applications, as the next wave of core methodology in statistics and machine learning will have to provide theoretical and computational tools to analyze graphs in order to support scientific progress in applied domains such as social sciences, biology, medicine, neuroscience, physics, finance, and economics.

While the field remains extremely heterogeneous and diverse, there are emerging signs of convergence, maturation, and increased awareness between the disparate disciplines. One noteworthy example, arising in studies on the spread of information, is that social media researchers are beginning to use problem-specific structure to infer between social influence, homophily, and external forces – subjects historically of intense interest amongst statisticians and social scientists. A second, more long-term example is the growing statistics literature expounding on topics popularized earlier within the physics community. Highly complex application domains, such as brain networks, are coming into the scope of the field.

**Goals:** The primary goal of the workshop is to become an inflection point in the maturation of social network and social media analysis, promoting greater technical sophistication and practical relevance. To accomplish this, we aim at bringing together researchers from applied disciplines such as sociology, economics,

## LOCATION:



medicine and biology, together with researchers from more theoretical disciplines such as mathematics and physics, within our community of statisticians and computer scientists.

The technical focus of the workshop is the statistical, methodological and computational issues that arise when modeling and analyzing large collections of data that are largely represented as static and dynamic graphs. As the different communities use diverse ideas and mathematical tools, we seek to foster cross-disciplinary collaborations and intellectual exchange. The communities identified above all have a long-standing interest in modeling networks, and while they approach the problem from different directions, their ultimate goals are very similar.

The NIPS community serves as the perfect middle ground to enable effective communication of both applied and methodological concerns. Earlier workshops that we organized at ICML 2006 (on "Statistical Network Analysis: Models, Issues and New Directions", in LNCS collection, vol. 4503.), and at NIPS 2008 (on "Analyzing Graphs: Theory and Applications"), were well attended, with standing audience at many talks. Since then, network modeling has grown to become a regular part of most if not all of the major conferences that are related to NIPS, and workshops held at NIPS 2009 and 2010 were large by the conference standards (100+ attendees) with overflowing attendance. This year, we are aware of a number of new compelling results and ongoing work, particularly in social media analysis, which reflect increasing sophistication and relevance of the field.

We would like to bring together a diverse set of researchers once again to assess progress and stimulate further debate, in an effort to support a continued, open, cross-disciplinary dialogue. We believe this effort will ultimately result in novel modeling approaches, and ultimately in the identification of diverse applications and open problems that may serve as guidance for future research directions.

## WEBSITE:

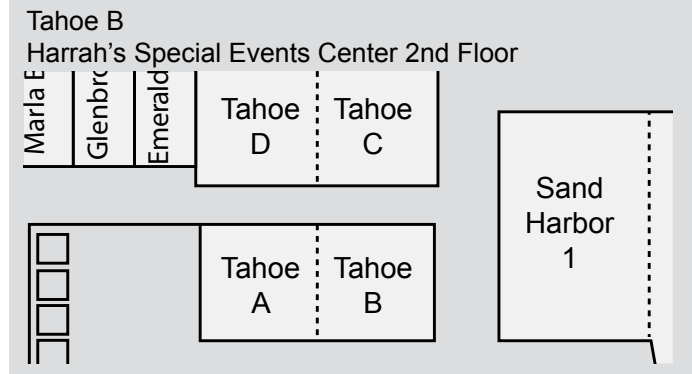
<http://snap.stanford.edu/social2012/>

# Spectral Algorithms for Latent Variable Models

## ORGANIZERS:

Ankur Parikh	Carnegie Mellon University
Le Song	Georgia Tech
Eric Xing	Carnegie Mellon University

## LOCATION:



## ABSTRACT:

Recently, linear algebra techniques have given a fundamentally different perspective for learning and inference in latent variable models. Exploiting the underlying spectral properties of the model parameters has led to fast, provably consistent methods for structure and parameter learning that stand in contrast to previous approaches, such as Expectation Maximization, which suffer from local optima and slow convergence. Furthermore, these techniques have given insight into the nature of latent variable models.

In this workshop, via a mix of invited talks, contributed posters, and discussion, we seek to explore the theoretical and applied aspects of spectral methods including the following major themes:

- (1) How can spectral techniques help us develop fast and local minima free solutions to real world problems involving latent variables in natural language processing, dynamical systems, computer vision etc. where existing methods such as Expectation Maximization are unsatisfactory?
- (2) How can these approaches lead to a deeper understanding and interpretation of the complexity of latent variable models?

## WEBSITE:

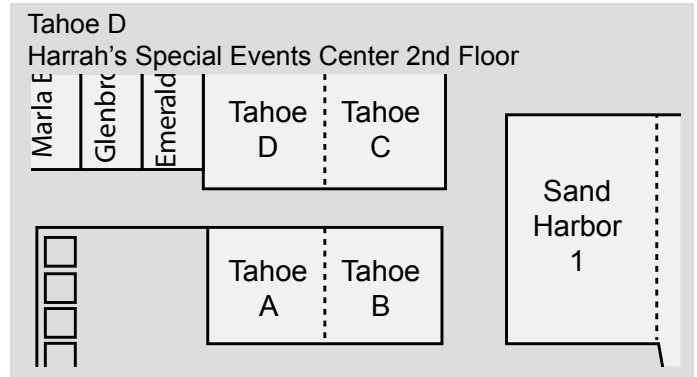
<http://www.cs.cmu.edu/~apparikh/nips2012spectral/main.html>

# xLiTe: Cross-Lingual Technologies

## ORGANIZERS:

Achim Rettinger	Karlsruhe Institute of Technology
Marko Grobelnik	Jozef Stefan Institute
Blaz Fortuna	Jozef Stefan Institute
Xavier Carreras	Technical University of Catalonia
Juanzi Li	Tsinghua University

## LOCATION:



## ABSTRACT:

### Workshop Motivation:

Automatic text understanding has been an unsolved research problem for many years. This partially results from the dynamic and diverging nature of human languages, which ultimately results in many different varieties of natural language. This variations range from the individual level, to regional and social dialects, and up to seemingly separate languages and language families.

However, in recent years there have been considerable achievements in data driven approaches to computational linguistics exploiting the redundancy in the encoded information and the structures used. Those approaches are mostly not language specific or can even exploit redundancies across languages.

This progress in cross-lingual technologies is largely due to the increased availability of multilingual data in the form of static repositories or streams of documents. In addition parallel and comparable corpora like Wikipedia are easily available and constantly updated. Finally, cross-lingual knowledge bases like DBpedia can be used as an Interlingua to connect structured information across languages. This helps at scaling the traditionally monolingual tasks, such as information retrieval and intelligent information access, to multilingual and cross-lingual applications.

From the application side, there is a clear need for such cross-lingual technology and services, as a) there is a huge disparity between the mean size of languages and the median size. It turns out that 389 (or nearly 6%) of the world's languages have at least one million speakers and account for 94% of the world's population. By contrast, the remaining 94% of languages are spoken by only 6% of the world's people. And b) in many areas like in the EU member states, 56% of the citizens are able to hold a conversation in one language apart from their mother tongue.

Available systems on the market are typically focused on multilingual tasks, such as machine translation, and don't deal with cross-linguality. A good example is one of the most popular news aggregators, namely Google News that collects news isolated per individual language. The ability to cross the border of a particular language would help many users to consume the breadth of news reporting by joining information in their mother tongue with information from the rest of the world.

### Workshop Objectives:

The XLT workshop is aimed at techniques, which strive for flexibility making them applicable across languages and language varieties with less manual effort and manual labeled training data. Such approaches might also be beneficial for solving the pressing task of analyzing the continuously evolving natural language varieties that are not well formed. Such data typically originates from social media, like text messages, forum posts or tweets and often is highly domain dependent.

## WEBSITE:

<http://km.aifb.kit.edu/ws/xlite/>



# **SATURDAY WORKSHOPS**



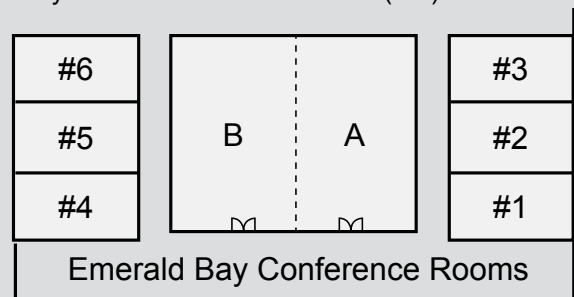
# Algebraic Topology and Machine Learning

## ORGANIZERS:

Sivaraman Balakrishnan	Carnegie Mellon University
Alessandro Rinaldo	Carnegie Mellon University
Donald Sheehy	INRIA Saclay-Ile-de-France
Aarti Singh	Carnegie Mellon University
Larry Wasserman	Carnegie Mellon University

## LOCATION:

Emerald Bay 6  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Topological methods and machine learning have long enjoyed fruitful interactions as evidenced by popular algorithms like ISOMAP, LLE and Laplacian Eigenmaps which have been borne out of studying point cloud data through the lens of topology/geometry. More recently several researchers have been attempting to understand the algebraic topological properties of data. Algebraic topology is a branch of mathematics which uses tools from abstract algebra to study and classify topological spaces. The machine learning community thus far has focussed almost exclusively on clustering as the main tool for unsupervised data analysis. Clustering however only scratches the surface, and algebraic topological methods aim at extracting much richer topological information from data.

The goals of our workshop are:

1. To draw the attention of machine learning researchers to a rich and emerging source of interesting and challenging problems.
2. To identify problems of interest to both topologists and machine learning researchers and areas of potential collaboration.
3. To discuss practical methods for implementing topological data analysis methods.
4. To discuss applications of topological data analysis to scientific problems.

We will also target submissions in a variety of areas, at the intersection of algebraic topology and learning, that have witnessed recent activity. Areas of focus for submissions include but are not limited to:

1. Robust geometric inference: New methods to understand the various topological properties of (possibly) noisy random samples, and also the applications of geometric inference techniques to other problems in machine learning (eg. supervised learning).
2. Applications of topological data analysis to new/existing areas: Topological data analysis has previously found application in a variety of areas including medical imaging and neuroscience, sensor networks, landmark-based shape data analyses, proteomics, microarray analysis and cellular biology.
3. Computational aspects of topological inference: Computing topological properties of data often involves building simplicial complexes and computing linear algebraic properties (eg. rank) of matrices associated with these complexes. These methods are typically computationally intensive, and new methods are being continually developed to help scale topological methods to larger datasets.
4. Statistical approaches to topological data analysis: Some recent work in computational topology has focussed on making statistical sense of topological properties. For instance, studying the topological properties of bootstrap samples and using this to associate topological properties of random data with confidence statements typical of statistical hypothesis testing.

## WEBSITE:

<https://sites.google.com/site/nips2012topology/>

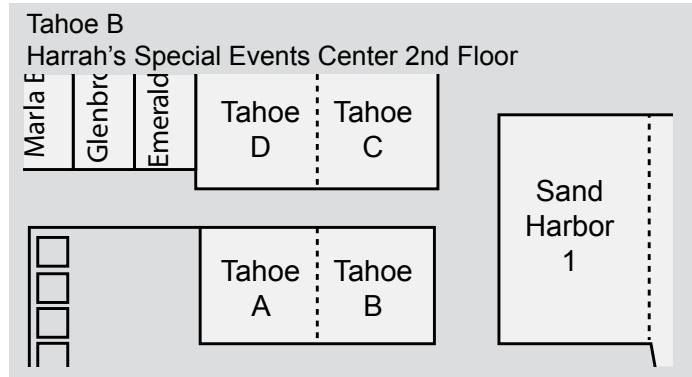
**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

# Bayesian Nonparametric Models For Reliable Planning And Decision-Making Under Uncertainty

## ORGANIZERS:

Jonathan How	Massachusetts Institute of Technology
Lawrence Carin	Duke
John Fisher	Massachusetts Institute of Technology
Michael Jordan	UC Berkeley
Alborz Geramifard	Massachusetts Institute of Technology

## LOCATION:



## ABSTRACT:

The ability to autonomously plan a course of action to reach a desired goal in the presence of uncertainty is critical for the success of autonomous robotic systems. Autonomous planning typically involves selecting actions that maximize the mission objectives given the available information, such as models of the agent dynamics, environment, available resources, and mission constraints. However, such models are typically only approximate, and can rapidly become obsolete, thereby degrading the planner performance. Classical approaches to address this problem typically assume that the environment has a certain structure that can be captured by a parametric model that can be updated online. However, finding the right parameterization a priori for complex and uncertain domains is challenging because substantial domain knowledge is required. An alternative approach is to let the data provide the insight on the parameterization. This approach leads to Bayesian Nonparametric models (BNPMs), which is a powerful framework for reasoning about objects and relations in settings in which these objects and relations are not predefined. This feature is particularly attractive for missions, such as long-term persistent sensing, for which it is virtually impossible to specify the size of the model and the number of parameters a priori. In such scenarios, BNPMs are especially well suited for integrating data from multiple sensors, and choosing the appropriate model size based on the observed data.

Gaussian processes (GPs) are an example of a widely used BNPM for regression and clustering problems. However, GPs are only one class of a rapidly growing number of BNPMs that capture, for example, hybrid system dynamics with an unknown number of modes or possibly shared features. While GPs have been used with some success in planning and decision-making applications, the use of other types of BNPMs for these scenarios is much less widespread. This is unfortunate because alternate BNPM techniques, such as Dirichlet Processes (DP), Beta Processes (BP), and their hierarchical variants, are potentially much better suited for discrete classification, clustering, and high level planning that involve discrete decisions.

Thus the purpose of this workshop is to provide a mechanism for the statistical machine learning and autonomous decision-making communities to discuss recent results in BNPM and inference techniques with the goal of investigating how the BNPMs can be used to improve the reliability and performance of autonomous planning and decision making systems in a data-rich world. This will be accomplished by bringing together leading experts and practicing engineers in both fields. The workshop will present a mix of invited sessions, contributed talks, poster session by leading experts and active researchers in BNPMs, robotics, and decision making and planning. Furthermore, the workshop schedule is designed to allow for plenty of mingle and question time that is expected to be conducive to merging of ideas and initiating collaborations.

## WEBSITE:

<http://acl.mit.edu/bnpm>

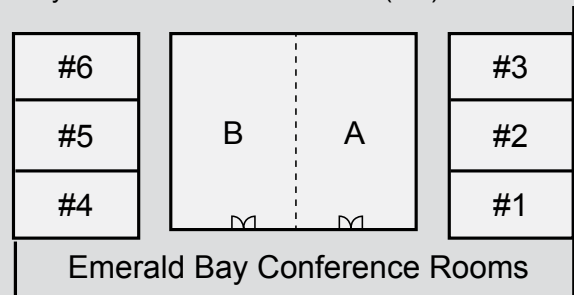
# Big Learning : Algorithms, Systems, and Tools

## ORGANIZERS:

Sameer Singh	UMass, Amherst
John Duchi	UC Berkeley
Yucheng Low	Carnegie Mellon University
Joseph Gonzalez	Carnegie Mellon University

## LOCATION:

Emerald Bay A  
Harveys Convention Center Floor (CC)



## ABSTRACT:

This workshop will address algorithms, systems, and real-world problem domains related to large-scale machine learning (“Big Learning”). With active research spanning machine learning, databases, parallel and distributed systems, parallel architectures, programming languages and abstractions, and even the sciences, Big Learning has attracted intense interest. This workshop will bring together experts across these diverse communities to discuss recent progress, share tools and software, identify pressing new challenges, and to exchange new ideas. Topics of interest include (but are not limited to):

- Big Data: Methods for managing large, unstructured, and/or streaming data; cleaning, visualization, interactive platforms for data understanding and interpretation; sketching and summarization techniques; sources of large datasets.
- Models & Algorithms: Machine learning algorithms for parallel, distributed, GPGPUs, or other novel architectures; theoretical analysis; distributed online algorithms; implementation and experimental evaluation; methods for distributed fault tolerance.
- Applications of Big Learning: Practical application studies and challenges of real-world system building; insights on end-users, common data characteristics (stream or batch); trade-offs between labeling strategies (e.g., curated or crowd-sourced).
- Tools, Software & Systems: Languages and libraries for large-scale parallel or distributed learning which leverage cloud computing, scalable storage (e.g. RDBMs, NoSQL, graph databases), and/or specialized hardware.

## WEBSITE:

<http://www.biglearn.org>

**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

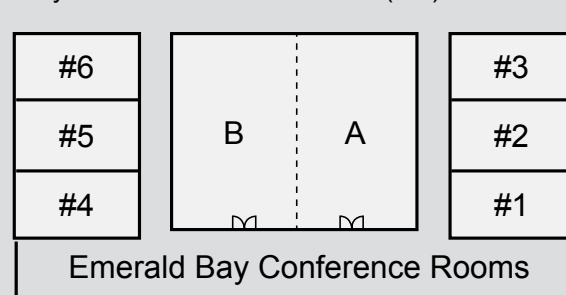
# Confluence between Kernel Methods and Graphical Models

## ORGANIZERS:

Le Song	Georgia Tech
Arthur Gretton	Gatsby Unit, UCL
Alexander Smola	Google

## LOCATION:

Emerald Bay 2  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Kernel methods and graphical models are two important families of techniques for machine learning. Our community has witnessed many major but separate advances in the theory and applications of both subfields. For kernel methods, the advances include kernels on structured data, Hilbert-space embeddings of distributions, and applications of kernel methods to multiple kernel learning, transfer learning, and multi-task learning. For graphical models, the advances include variational inference, nonparametric Bayes techniques, and applications of graphical models to topic modeling, computational biology and social network problems.

This workshop addresses two main research questions: first, how may kernel methods be used to address difficult learning problems for graphical models, such as inference for multi-modal continuous distributions on many variables, and dealing with non-conjugate priors? And second, how might kernel methods be advanced by bringing in concepts from graphical models, for instance by incorporating sophisticated conditional independence structures, latent variables, and prior information?

Kernel algorithms have traditionally had the advantage of being solved via convex optimization or eigenproblems, and having strong statistical guarantees on convergence. The graphical model literature has focused on modelling complex dependence structures in a flexible way, although approximations may be required to make inference tractable. Can we develop a new set of methods which blend these strengths?

There has recently been a number of publications combining kernel and graphical model techniques, including kernel hidden Markov models [SBSGS08], kernel belief propagation [SGBLG11], kernel Bayes rule [FSG11], kernel topic models [HSHG12], kernel variational inference [GHB12], kernel herding as Bayesian quadrature [HD12], kernel beta processes [RWDC08], a connection between kernel k-means and Bayesian nonparametrics [KJ12] and kernel determinantal point processes for recommendations [KT12]. Each of these results deals with different inference tasks, and makes use of a range of RKHS properties. We propose this workshop so as to “connect the dots” and develop a unified toolkit to address a broad range of learning problems, to the mutual benefit of researchers in kernels and graphical models. The goals of the workshop are thus twofold: first, to provide an accessible review and synthesis of recent results combining graphical models and kernels. Second, to provide a discussion forum for open problems and technical challenges.

See website for bibliographies

## WEBSITE:

<https://sites.google.com/site/kernelgraphical>

**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

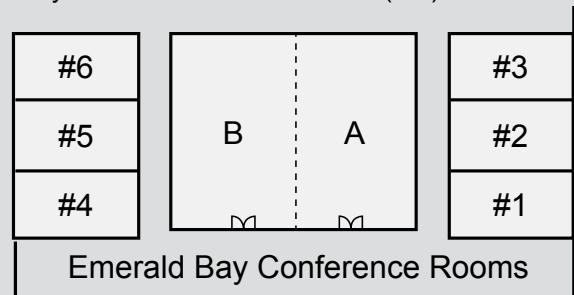
# Deep Learning and Unsupervised Feature Learning

## ORGANIZERS:

Yoshua Bengio	University of Montreal
James Bergstra	University of Waterloo
Quoc V. Le	Stanford University

## LOCATION:

Emerald Bay B  
Harveys Convention Center Floor (CC)



## ABSTRACT:

Machine learning algorithms are very sensitive to the representations chosen for the data so it is desirable to improve learning algorithms that can discover good representations, good features, or good explanatory latent variables. Both supervised and unsupervised learning algorithms have been proposed for this purpose, and they can be combined in semi-supervised setups in order to take advantage of vast quantities of unlabeled data. Deep learning algorithms have multiple levels of representation and the number of levels can be selected based on the available data. Great progress has been made in recent years in algorithms, their analysis, and their application both in academic benchmarks and large-scale industrial settings (such as machine vision/object recognition and NLP, including speech recognition). Many interesting open problems also remain, which should stimulate lively discussions among the participants.

## WEBSITE:

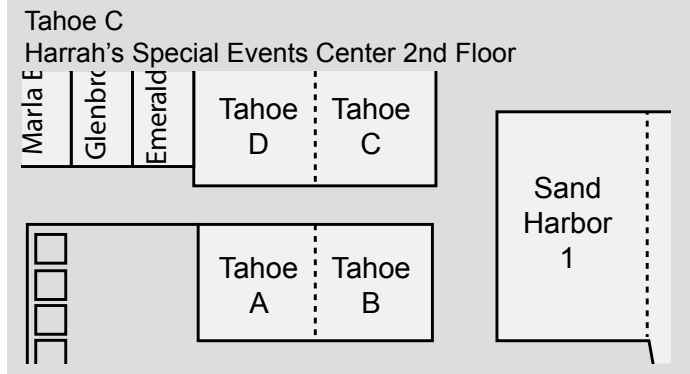
<https://sites.google.com/site/deeplearningnips2012/>

# Log-Linear Models

## ORGANIZERS:

Dimitri Kanevsky	IBM Research Center
Tony Jebara	Columbia University
Li Deng	Microsoft Research
Stephen Wright	University of Wisconsin-Madison
Georg Heigold	Google
Avishy Carmi	NTU

## LOCATION:



## ABSTRACT:

Exponential functions are core mathematical constructs that are the key to many important applications, including speech recognition, pattern-search and logistic regression problems in statistics, machine translation, and natural language processing. Exponential functions are found in exponential families, log-linear models, conditional random fields (CRF), entropy functions, neural networks involving sigmoid and soft max functions, and Kalman filter or MMIE training of hidden Markov models. Many techniques have been developed in pattern recognition to construct formulations from exponential expressions and to optimize such functions, including growth transforms, EM, EBW, Rprop, bounds for log-linear models, large-margin formulations, and regularization. Optimization of log-linear models also provides important algorithmic tools for machine learning applications (including deep learning), leading to new research in such topics as stochastic gradient methods, sparse / regularized optimization methods, enhanced first-order methods, coordinate descent, and approximate second-order methods. Specific recent advances relevant to log-linear modeling include the following.

- Effective optimization approaches, including stochastic gradient and Hessian-free methods.
- Efficient algorithms for regularized optimization problems.
- Bounds for log-linear models and recent convergence results
- Recognition of modeling equivalences across different areas, such as the equivalence between Gaussian and log-linear models/HMM and HCRF, and the equivalence between transfer entropy and Granger causality for Gaussian parameters.

Though exponential functions and log-linear models are well established, research activity remains intense, due to the central importance of the area in front-line applications and the rapid expanding size of the data sets to be processed. Fundamental work is needed to transfer algorithmic ideas across different contexts and explore synergies between them, to assimilate the influx of ideas from optimization, to assemble better combinations of algorithmic elements for tackling such key tasks as deep learning, and to explore such key issues as parameter tuning.

The workshop will bring together researchers from the many fields that formulate, use, analyze, and optimize log-linear models, with a view to exposing and studying the issues discussed above.

## WEBSITE:

<https://sites.google.com/site/nips12logmodels/home>



**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

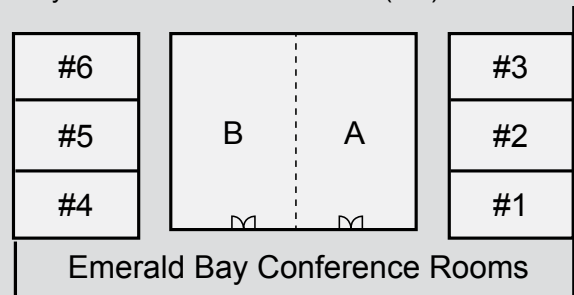
# Machine Learning Approaches to Mobile Context Awareness

## ORGANIZERS:

Kat Ellis	UC San Diego
Gert Lanckriet	UC San Diego
Tommi Jaakkola	Massachusetts Institute of Technology
Lenny Grokop	Qualcomm

## LOCATION:

**Emerald Bay 3**  
**Harveys Convention Center Floor (CC)**



## ABSTRACT:

The ubiquity of mobile phones, packed with sensors such as accelerometers, gyroscopes, light and proximity sensors, Bluetooth and WiFi radios, GPS radios, microphones, etc., has brought increased attention to the field of mobile context awareness. This field examines problems relating to inferring some aspect of a user's behavior, such as their activity, mood, interruptibility, situation, etc., using mobile sensors. There is a wide range of applications for context-aware devices. In the healthcare industry, for example, such devices could provide support for cognitively impaired people, provide health-care professionals with simple ways of monitoring patient activity levels during rehabilitation, and perform long-term health and fitness monitoring. In the transportation industry they could be used to predict and redirect traffic flow or to provide telematics for auto-insurers. Context awareness in smartphones can aid in automating functionality such as redirecting calls to voicemail when the user is uninterruptible, automatically updating status on social networks, etc., and can be used to provide personalized recommendations.

Existing work in mobile context-awareness has predominantly come from researchers in the human-computer interaction community. There the focus has been on building custom sensor/hardware solutions to perform social science experiments or solve application-specific problems. The goal of this workshop is to bring the challenging inferential problems of mobile context awareness to the attention of the machine learning community. We believe these problems are fundamentally solvable. We seek to get this community excited about these problems, encourage collaboration between people with different backgrounds, explore how to integrate research efforts, and discuss where future work needs to be done. We are looking for participation both from

individuals with machine learning backgrounds who may or may not have attacked context awareness problems before, and individuals with application-specific backgrounds. Although the dominant mobile sensing platform these days is the smartphone, we also welcome contributions that work with data from a variety of body-worn sensors including standalone accelerometers, GPS, microphones, EEG, ECG, etc., and custom hardware platforms that combine multiple sensors. We are particularly interested in contributions that deal with inferring context by fusing information from different sensor sources.

In particular, we would like the workshop to address the following topics:

- (1) What is the best way to combine heterogeneous data from multiple sensors? Is contextual information encoded in specific correlation patterns, or is there one sensor that "says it all" for each context, and can we learn this automatically? How do we model and analyze correlations between heterogeneous data?
- (2) Feature extraction: what are the features that best characterize these new sensor streams for analysis and learning? In video and speech processing, such features have emerged over the years and are now commonly accepted – are there certain features best suited for accelerometer, audio environment, and GPS data streams? Can we learn them automatically?
- (3) A major part of this workshop will be dedicated to the discussion of data. The community has a great need for a shared public dataset that will allow researchers to compare algorithms and improve collaboration. In our panel discussion we will discuss issues such as creating a central data repository, common data collection apps, and unique issues with context-awareness data.

## WEBSITE:

<https://sites.google.com/site/nips2012contextawareworkshop/>

**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

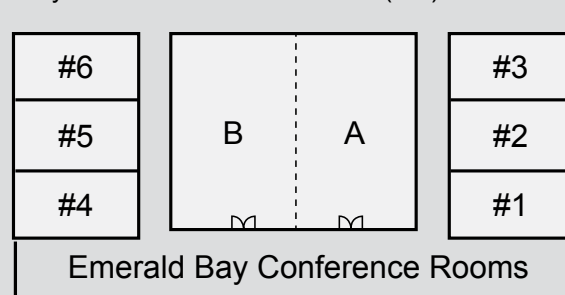
# MLINI - 2nd NIPS Workshop on Machine Learning and Interpretation in Neuroimaging (2-day)

## ORGANIZERS:

Georg Langs	Medical University of Vienna
Irina Rish	IBM Research
Guillermo Cecchi	IBM Research
Brian Murphy	Carnegie Mellon University
Bjoern Menze	ETH Zurich
Kai-min Chang	Carnegie Mellon University
Moritz Grosse-Wentrup	Max Planck Institute

## LOCATION:

Emerald Bay 5  
Harveys Convention Center Floor (CC)



## ABSTRACT:

We propose a two day workshop on the topic of machine learning approaches in neuroscience and neuroimaging. We believe that both machine learning and neuroimaging can learn from each other as the two communities overlap and enter an intense exchange of ideas and research questions. Methodological developments in machine learning spurn novel paradigms in neuroimaging, neuroscience motivates methodological advances in computational analysis. In this context many controversies and open questions exist. The goal of the workshop is to pinpoint these issues, sketch future directions, and tackle open questions in the light of novel methodology.

The first workshop of this series at NIPS 2011 built upon earlier events in 2006 and 2008. Last year's workshop included many invited speakers, and was centered around two panel discussions, during which 2 questions were discussed: the interpretability of machine learning findings, and the shift of paradigms in the neuroscience community. The discussion was inspiring, and made clear, that there is a tremendous amount the two communities can learn from each other benefiting from communication across the disciplines.

The aim of the workshop is to offer a forum for the overlap of these communities. Besides interpretation, and the shift of paradigms, many open questions remain. Among them:

- How suitable are MVPA and inference methods for brain mapping?
- How can we assess the specificity and sensitivity?
- What is the role of decoding vs. embedded or separate feature selection?
- How can we use these approaches for a flexible and useful representation of neuroimaging data?
- What can we accomplish with generative vs. discriminative modelling?
- Can and should the Machine Learning community provide a standard repertoire of methods for the Neuroimaging community to use (e.g. in choosing a classifier)?

For a detailed background, please visit the website.

## WEBSITE:

<https://sites.google.com/site/nipsmlini2012/>

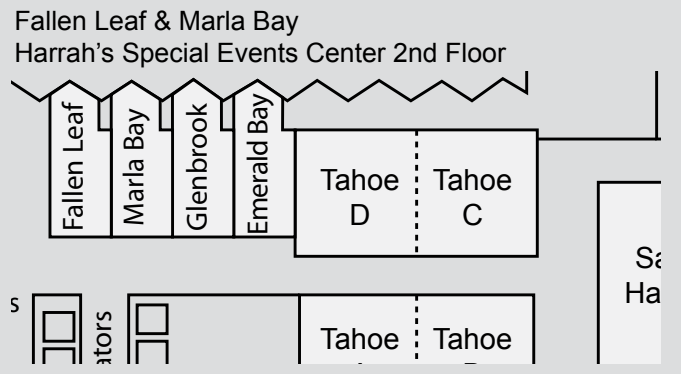
**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

# Optimization for Machine Learning

## ORGANIZERS:

Suvrit Sra	Max Planck Institute
Alekh Agarwal	Microsoft Research

## LOCATION:



## ABSTRACT:

Optimization lies at the heart of ML algorithms. Sometimes, classical textbook algorithms suffice, but the majority problems require tailored methods that are based on a deeper understanding of the ML requirements. ML applications and researchers are driving some of the most cutting-edge developments in optimization today. The intimate relation of optimization with ML is the key motivation for our workshop, which aims to foster discussion, discovery, and dissemination of the state-of-the-art in optimization as relevant to machine learning.

Much interest has focused recently on stochastic methods, which can be used in an online setting and in settings where data sets are extremely large and high accuracy is not required. Many aspects of stochastic gradient remain to be explored, for example, different algorithmic variants, customizing to the data set structure, convergence analysis, sampling techniques, software, choice of regularization and tradeoff parameters, distributed and parallel computation. The need for an up-to-date analysis of algorithms for nonconvex problems remains an important practical issue, whose importance becomes even more pronounced as ML tackles more and more complex mathematical models.

Finally, we do not wish to ignore the not particularly large scale setting, where one does have time to wield substantial computational resources. In this setting, high-accuracy solutions and deep understanding of the lessons contained in the data are needed. Examples valuable to MLers may be exploration of genetic and environmental data to identify risk factors for disease; or problems dealing with setups where the amount of observed data is not huge, but the mathematical model is complex.

## WEBSITE:

<http://opt.kyb.tuebingen.mpg.de/>

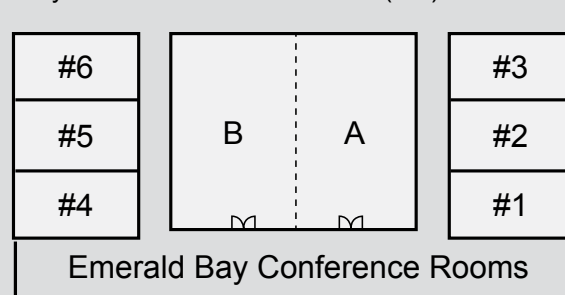
# Personalizing Education with Machine Learning

## ORGANIZERS:

Michael Mozer	University of Colorado
javier movellan	UC San Diego
Robert Lindsey	University of Colorado
Jacob Whitehill	UC San Diego

## LOCATION:

Emerald Bay 1  
Harveys Convention Center Floor (CC)



## ABSTRACT:

The field of education has the potential to be transformed by the internet and intelligent computer systems. Evidence for the first stage of this transformation is abundant, from the Stanford online AI and Machine Learning courses to web sites such as Kahn Academy that offer on line lessons and drills. However, the delivery of instruction via web-connected devices is merely a precondition for what may become an even more fundamental transformation: the personalization of education.

In traditional classroom settings, teachers must divide their attention and time among many students and hence have limited ability to observe and customize instruction to individuals. Even in one-on-one tutoring sessions, teachers rely on intuition and experience to choose the material and style of instruction that they believe would provide the greatest benefit given the student's current state of understanding.

In order both to assist human teachers in traditional classroom environments and to improve automated tutoring systems to match the capabilities of expert human tutors, one would like to develop formal approaches that can:

- exploit subtle aspects of a student's behavior---such as facial expressions, fixation sequences, response latencies, and errors---to make explicit inferences about the student's latent state of knowledge and understanding;
- leverage the latent state to design teaching policies and methodologies that will optimize the student's knowledge acquisition, retention, and understanding; and

- personalize instruction by providing material and interaction suited to the capabilities and preferences of the student.

Machine learning provides a rich set of tools, extending classical psychometric approaches, for data-driven latent state inference, policy optimization, and personalization. Years ago, it would have been difficult to obtain enough data for a machine learning approach. However, online interactions with students have become commonplace, and these interactions yield a wealth of data. The data to be mined go beyond what is typed: Cameras and microphones are ubiquitous on portable devices, allowing for the exploitation of subtle video and audio cues. Because web-based instruction offers data from a potentially vast collection of diverse learners, the population of learners should serve useful in drawing inferences about individual learners.

Mining the vast datasets on teaching and learning that emerge over the coming years may both yield important insights into effective teaching strategies and also deliver practical tools to assist both human and automated teachers.

The goal of this workshop is to bring together researchers in machine learning, data mining, and computational statistics with researchers in education, psychometrics, intelligent tutoring systems, and designers of web-based instructional software. Although a relatively young journal and conference on educational data mining has been established ([educationaldatamining.org](http://educationaldatamining.org)), the field hasn't had as much contact with machine learning theoreticians as one would like.

## WEBSITE:

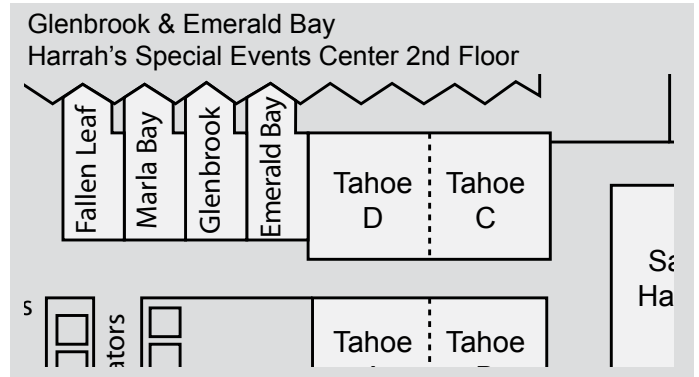
<http://www.cs.colorado.edu/~mozer/Admin/PersonalizingEducationNIPS2012/>

# Perturbations, Optimization, and Statistics

## ORGANIZERS:

Tamir Hazan	TTI Chicago
George Papandreou	UCLA
Danny Tarlow	University of Toronto

## LOCATION:



## ABSTRACT:

In nearly all machine learning tasks, we expect there to be randomness, or noise, in the data we observe and in the relationships encoded by the model. Usually, this noise is considered undesirable, and we would eliminate it if possible. However, there is an emerging body of work on perturbation methods, showing the benefits of explicitly adding noise into the modeling, learning, and inference pipelines. This workshop will bring together the growing community of researchers interested in different aspects of this area, and will broaden our understanding of why and how perturbation methods can be useful.

More generally, perturbation methods usually provide efficient and principled ways to reason about the neighborhood of possible outcomes when trying to make the best decision. For example, some might want to arrive at the best outcome that is robust to small changes in model parameters. Others might want to find the best choice while compensating for their lack of knowledge by averaging over the different outcomes. Recently, several works influenced by diverse fields of research such as statistics, optimization, machine learning, and theoretical computer science, use perturbation methods in similar ways. The goal of this workshop is to explore different techniques in perturbation methods and their consequences on computation, statistics and optimization. We shall specifically be interested in understanding the following issues:

- Statistical Modeling: What types of statistical models can be defined for structured prediction? How can random perturbations be used to relate computation and statistics?
- Efficient Sampling: What are the computational properties that allow efficient and unbiased sampling? How do perturbations control the geometry of such models and how can we construct sampling methods for these families?

- Approximate Inference: What are the computational and statistical requirements from inference? How can the maximum of random perturbations be used to measure the uncertainty of a system?
- Learning: How can we probabilistically learn model parameters from training data using random perturbations? What are the connections with max-margin and conditional random fields techniques?
- Theory: How does the maximum of a random process relate to its complexity? What are the statistical and computational properties it describes in Gaussian free fields over graphs?
- Pseudo-sampling: How do dynamical systems encode randomness? To what extent do perturbations direct us to the "pseudo-randomness" of its underlying dynamics?
- Robust classification: How can classifiers be learned in a robust way, and how can support vector machines be realized in this context? What are the relations between adversarial perturbations and regularizations and what are their extensions to structured predictions?
- Robust reconstructions: How can information be robustly encoded? In what ways can learning be improved by perturbing the input measurements?
- Adversarial Uncertainty: How can structured prediction be performed in zero-sum game setting? What are the computational qualities of such solutions, and do Nash-equilibria exist in these cases?

**Target Audience:** The workshop should appeal to NIPS attendees interested in both theoretical aspects such as Bayesian modeling, Monte Carlo sampling, optimization, inference, and learning, as well as practical applications in computer vision and language modeling.

## WEBSITE:

<http://www.stat.ucla.edu/~gpapan/pos12>

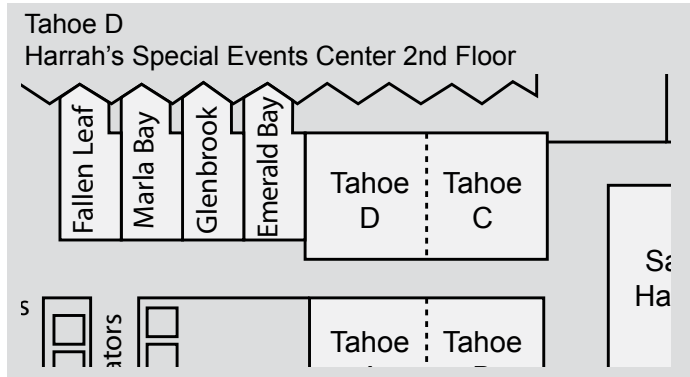
**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

# Probabilistic Numerics

## ORGANIZERS:

Philipp Hennig	Max Planck Institute
John Cunningham	University of Cambridge
Michael Osborne	University of Oxford

## LOCATION:



## ABSTRACT:

Traditionally, machine learning uses numerical algorithms as tools. However, many tasks in numerics can be viewed as learning problems. As examples:

- How can optimizers learn about the objective function, and how should they update their search direction?
- How should a quadrature method estimate an integral given observations of the integrand, and where should these methods put their evaluation nodes?
- Can approximate inference techniques be applied to numerical problems?

Many such issues can be seen as special cases of decision theory, active learning, or reinforcement learning.

We invite contribution of recent results in the development and analysis of numerical analysis methods based on probability theory. This includes, but is not limited to the areas of optimization, sampling, linear algebra, quadrature and the solution of differential equations.

## WEBSITE:

<http://www.probablistic-numerics.org>

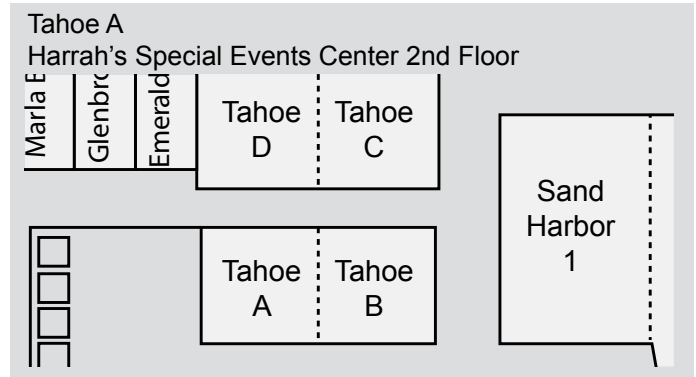


# Probabilistic Programming: Foundations and Applications (2-day)

## ORGANIZERS:

Vikash Mansinghka	MIT and Prior Knowledge, Inc
Dan Roy	University of Cambridge
Noah Goodman	Stanford University

## LOCATION:



## ABSTRACT:

Probabilistic models and algorithmic techniques for inference have become standard tools for interpreting data and building systems that learn from their experience. Growing out of an extensive body of work in machine learning, statistics, robotics, vision, artificial intelligence, neuroscience and cognitive science, rich probabilistic models and inference techniques have more recently spread to other branches of science and engineering, from astrophysics to climate science to marketing to web site personalization. This explosion is largely due to the development of probabilistic graphical models, which provide a formal lingua franca for modeling, and a common target for efficient inference algorithms.

However, even simple probabilistic models can require significant effort and specialized expertise to develop and use, frequently involving custom mathematics, algorithm design and software development. More innovative and useful models far outstrip the representational capacity of graphical models and their associated inference techniques. They are communicated using a mix of natural language, pseudo code, and formulas, often eliding crucial aspects such as fine-grained independence, abstraction and recursion, and are fit to data via special purpose, one-off inference algorithms.

PROBABILISTIC PROGRAMMING LANGUAGES aim to close this gap, going beyond graphical models in representational capacity while providing automatic probabilistic inference. Rather than marry statistics with graph theory, probabilistic programming marries Bayesian probability with universal computation. Instead of modeling joint distributions over a set of random variables, probabilistic programs model distributions over the execution histories of programs, including programs that analyze, transform and write other programs. Users specify a probabilistic model in its entirety (e.g., by writing code that generates a sample from the joint distribution)

and inference follows automatically given the specification. These languages provide the full power of modern programming languages for describing complex distributions, and can enable reuse of libraries of models, support interactive modeling and formal verification, and provide a much-needed abstraction barrier to foster generic, efficient inference in universal model classes.

We believe that the probabilistic programming language approach within AI has the potential to fundamentally change the way we understand, design, build, test and deploy probabilistic systems. This approach has seen growing interest within AI over the last 10 years, and builds on over 40 years of work in range of diverse fields including mathematical logic, theoretical computer science, programming languages, as well as machine learning, computational statistics, systems biology, probabilistic AI. However, considerable engineering challenges need to be solved before these techniques can be broadly adopted --- such as making inference efficient and producing robust, general-purpose software for probabilistic programming --- as do foundational questions about the complexity of inference and learning.

Our 2008 NIPS workshop helped to create the probabilistic programming community. For our 2012 workshop, we propose to:

- Assess and synthesize progress in probabilistic programming since 2008, including progress in languages like BLOG, Church, Figaro, ICL, Markov Logic, CSoft/Infer.NET and ProbLog
- Organize the growing community around key technical problems and benchmarks, to spur and systematize progress in the development and analysis of probabilistic programming systems
- Expose interested funding agencies to probabilistic programming research

## WEBSITE:

[http://probabilistic-programming.org/wiki/NIPS\\*2012\\_Workshop](http://probabilistic-programming.org/wiki/NIPS*2012_Workshop)



**SATURDAY WORKSHOP**  
**7:30AM - 6:30PM**

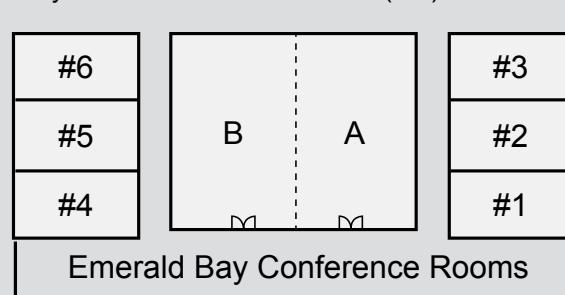
# Social Choice: Theory and Practice

## ORGANIZERS:

Behrouz Touri	University of Illinois
Olgica Milenkovic	University of Illinois
Faramarz Fekri	GTech

## LOCATION:

Emerald Bay 4  
Harveys Convention Center Floor (CC)



## ABSTRACT:

The aim of the workshop is to bring together a diverse research community from social choice theory, social sciences, and psychology on one side, and mathematics, statistics, decision and information theory on the other side. These fields traditionally have little interaction, despite the fact that a number of relevant practical problems in social sciences may be successfully addressed using techniques and tools developed in signal processing and information theory.

The organizers plan to create a forum for studying traditional and emerging experimental and computational problems in social choice theory through a multifaceted analytical lens.

The topics of interest include social decision making mechanisms, voting protocols and vote aggregation, voting and voter influence over social networks, and opinion dynamics modeling and monitoring. These topics lie at the heart of modern psychology, political sciences and sociology, and in addition, they have emerged in various incarnations in Internet applications, recommender systems and computer science in general.

The targeted group of participants will be selected experts in the field of social choice theory, decision theory, cognitive neuroscience, machine learning, information theory and statistics.

## WEBSITE:

<https://netfiles.uiuc.edu/touri1/www/nips/>

# LAKE TAHOE DESTINATION MAP

