# FRONTIERS IN COMPUTING + MATHEMATICAL SCIENCES

JANUARY 11 AND FEBRUARY 8 2021

Caltech

# FRONTIERS IN COMPUTING + MATHEMATICAL SCIENCES 2021

ALL TALKS ARE ONE HOUR LONG, AND WILL BE HELD ONLINE IN ZOOM.
PLEASE REFER TO E-ANNOUNCEMENTS FOR THE LINK, OR CONTACT SYDNEY@CALTECH.EDU.

## JANUARY 11

| | | |
|---|---|---|
| 8:55am | Introduction | Andrew Stuart<br>*Caltech* |
| 9:00am | Optimal Transport for Inverse Problems and the Implicit Regularization | Yunan Yang<br>*New York University* |
| 10:15am | Advancing Scalable, Provable Optimization Methods in Semidefinite & Polynomial Programs | Diego Cifuentes<br>*Massachusetts Institute of Technology* |
| 11:30am | Solving High-Dimensional PDEs, Controls, and Games with Deep Learning | Jiequn Han<br>*Princeton University* |
| 2:00pm | Proof Engineering Tools for a New Era | Talia Ringer<br>*University of Washington* |
| 3:15pm | Reliable Machine Learning in Feedback Systems | Sarah Dean<br>*University of California, Berkeley* |

# FRONTIERS IN COMPUTING + MATHEMATICAL SCIENCES 2021

ALL TALKS ARE ONE HOUR LONG, AND WILL BE HELD ONLINE IN ZOOM.
PLEASE REFER TO E-ANNOUNCEMENTS FOR THE LINK, OR CONTACT SYDNEY@CALTECH.EDU.

## FEBRUARY 8

| | | |
|---|---|---|
| 8:55am | Introduction | Andrew Stuart<br>*Caltech* |
| 9:00am | Demystifying the Efficiency of Reinforcement Learning: Two Recent Stories | Yuxin Chen<br>*Princeton University* |
| 10:15am | Bridging Learning and Decision Making | Dylan Foster<br>*Massachusetts Institute of Technology* |
| 11:30am | Integrating Machine Learning into Algorithm Design | Ellen Vitercik<br>*Carnegie Mellon University* |
| 2:00pm | Inferring Chaos and Emergent Low-Dimensionality in Living Dynamical Systems | William Gilpin<br>*Harvard Quantitative Biology Initiative* |

# Demystifying the Efficiency of Reinforcement Learning: Two Recent Stories


Yuxin Chen
*Princeton University*

Reinforcement learning (RL), which is frequently modeled as sequential learning and decision making in the face of uncertainty, is garnering growing interest in recent years due to its remarkable success in practice. In contemporary RL applications, it is increasingly more common to encounter environments with prohibitively large state and action space, thus imposing stringent requirements on the sample and computational efficiency of the RL algorithms in use. Despite the empirical success, however, the theoretical underpinnings for many popular RL algorithms remain highly inadequate even for the tabular setting.

In this talk, we present two vignettes regarding the effectiveness of RL algorithms. The first vignette demonstrates that a perturbed model-based RL approach is minimax optimal under a generative model, without suffering from a sample size barrier that was present in all past work. The second vignette covers policy optimization in reinforcement learning. On the one hand, we demonstrate that the popular softmax policy gradient method can take exponential time to converge; on the other hand, employing natural policy gradients and enforcing entropy regularization provably achieve fast global convergence. These results cover two distinctive RL paradigms, and might shed light on the efficacy of these algorithms in more complicated scenarios.

# Advancing Scalable, Provable Optimization Methods in Semidefinite & Polynomial Programs

Diego Cifuentes
*Massachusetts Institute of Technology*

Optimization is a broad area with ramifications in many disciplines, including machine learning, control theory, signal processing, robotics, computer vision, power systems, and quantum information. I will talk about some novel algorithmic and theoretical results in two broad classes of optimization problems. The first class of problems are semidefinite programs (SDP). I will present the first polynomial time guarantees for the Burer-Monteiro method, which is widely used for solving large scale SDPs. I will also discuss some general guarantees on the quality of SDP solutions for parameter estimation problems. The second class of problems I will consider are polynomial systems. I will introduce a novel technique for solving polynomial systems that, by taking advantage of graphical structure, is able to outperform existing techniques by orders of magnitude in suitable applications.

# Reliable Machine Learning in Feedback Systems

Sarah Dean
*University of California, Berkeley*

Machine learning techniques have been successful for processing complex information, and thus they have the potential to play an important role in data-driven decision-making and control. However, ensuring the reliability of these methods in feedback systems remains a challenge, since classic statistical and algorithmic guarantees do not always hold.

In this talk, I will provide rigorous guarantees of safety and agency in dynamical settings relevant to robotics and recommendation systems. I take a perspective based on reachability, to specify which parts of the state space the system avoids (safety) or can be driven to (agency). For data-driven control, we show finite-sample performance and safety guarantees which highlight relevant properties of the system to be controlled. For recommendation systems, we introduce a novel metric of agency and show that it can be efficiently computed. In closing, I discuss how the reachability perspective can be used to design social-digital systems with a variety of important values in mind.

# Bridging Learning and Decision Making

Dylan Foster
*Massachusetts Institute of Technology*

Machine learning is becoming widely used in decision making, in domains ranging from personalized medicine and mobile health to online education and recommendation systems. While (supervised) machine learning traditionally excels at prediction problems, decision making requires answering questions that are counterfactual in nature, and ignoring this mismatch leads to unreliable decisions. As a consequence, our understanding of the algorithmic foundations for data-driven decision making is limited, and efficient algorithms are typically developed on an ad hoc basis. Can we bridge this gap and make decision making as easy as machine learning?

Focusing on the contextual bandit, a core problem in data-driven decision making, we bridge the gap by providing the first optimal and efficient reduction to supervised machine learning. The algorithm allows users to seamlessly apply off-the-shelf supervised learning models and methods to make decisions on the fly, and has been implemented in widely-used, industry-standard tools for decision making.

Our results advance a broader program to develop a universal algorithm design paradigm for data-driven decision making. I will close the talk by discussing challenges and opportunities in building such a framework, including efforts to extend our developments to difficult reinforcement learning problems in large state spaces.

# Inferring Chaos and Emergent Low-Dimensionality in Living Dynamical Systems

William Gilpin
*Harvard Quantitative Biology Initiative*

Dynamical systems theory provides a rich set of tools for inferring underlying mathematical structure from partial observations of complex systems, yet translating these insights to real-world biological datasets remains challenging. In this talk, I will overview my recent work at the intersection of nonlinear dynamics, chaos, and biology. I will first focus on my recent work developing physics-informed machine learning algorithms that extract dynamical models directly from raw experimental data. I will present a general technique for discovering strange attractors within diverse biological time series, including gene expression, patient electrocardiograms, animal trackers, and neural spiking. Next, I will describe my work on biological fluid dynamics, and the discovery of a beautiful vortex array created as many invertebrates swim—which enables a novel feeding strategy based on chaotic mixing of the local microenvironment. I will relate this work to broader questions at the intersection of nonlinear dynamics and organismal behavior. I will discuss how these insights open up several exciting new avenues at the intersection of dynamical systems theory, systems biology, and machine learning.

# Solving High-Dimensional PDEs, Controls, and Games with Deep Learning

Jiequn Han
*Princeton University*

Developing algorithms for solving high-dimensional partial differential equations, controls, and games has been an exceedingly difficult task for a long time, due to the notorious "curse of dimensionality". In this talk, we introduce a family of deep learning-based algorithms for these problems. The algorithms exploit the mathematical structure of problems and utilize deep neural networks as efficient approximations to high-dimensional functions. Numerical results of a variety of examples demonstrate the efficiency and accuracy of the proposed algorithms in high-dimensions. This opens up new possibilities in economics, engineering, and physics, by considering all participating agents, assets, resources, or particles together at the same time, instead of making ad hoc assumptions on their inter-relationships.

# Proof Engineering Tools for a New Era

Talia Ringer
*University of Washington*

Interactive theorem provers make it possible to prove that a program satisfies a specification. This provides a high degree of certainty that the program is trustworthy. The last two decades have marked a new era of verification of large and critical systems using interactive theorem provers. Still, the costs of developing these verified systems are high, and the costs of maintaining them even higher. These costs are addressed by *proof engineering*: technologies that make it easier to develop and maintain verified systems.

This talk will present some of the key challenges that proof engineering addresses, focusing in particular on my work on *proof repair*. In contrast with traditional proof automation, proof repair views proofs as fluid entities that must evolve alongside the programs whose correctness they prove. My work on proof repair uses a combination of semantic differencing and program transformations on proof terms to adapt proofs to breaking changes. I have implemented these techniques in a flexible proof repair tool called PUMPKIN PATCH. PUMPKIN PATCH has already been used to support proof engineering benchmarks, ease development with dependent types, and simplify a mechanized verification of the TLS Handshake protocol.

# Integrating Machine Learning Into Algorithm Design

Ellen Vitercik
*Carnegie Mellon University*

An important property of those algorithms that are typically used in practice is broad applicability—the ability to solve problems across diverse domains. However, the default, out-of-the-box performance of these algorithms can be unsatisfactory, with slow runtime, poor solution quality, and even negative long-term social ramifications. In practice, there is often ample data available about the types of problems an algorithm will be run on, data that can potentially be harnessed to fine-tune the algorithm's performance. We therefore need principled approaches for using this data to obtain strong application-specific performance guarantees.

In this talk, I will give an overview of my research that provides practical methods built on firm theoretical foundations for incorporating machine learning and optimization into the process of algorithm design, selection, and configuration. I will describe my contributions across several diverse domains, including integer programming, clustering, mechanism design, and computational biology. As I will demonstrate, these seemingly disparate areas are connected by overarching structure which implies broadly-applicable guarantees.

# Optimal Transport for Inverse Problems and the Implicit Regularization

Yunan Yang
*New York University*

Optimal transport has been one interesting topic of mathematical analysis since Monge (1781). The problem's close connections with differential geometry and kinetic descriptions were discovered within the past century, and the seminal work of Kantorovich (1942) showed its power to solve real-world problems. Recently, we proposed the quadratic Wasserstein distance from optimal transport theory for inverse problems, tackling the classical least-squares method's longstanding difficulties such as nonconvexity and noise sensitivity. The work was soon adopted in the oil industry. As we advance, we discover that the advantage of changing the data misfit is more general in a broader class of data-fitting problems by examining the preconditioning and "implicit" regularization effects of different mathematical metrics as the objective function in optimization, as the likelihood function in Bayesian inference, and as the measure of residual in numerical solution to PDEs.