THE FIELDS INSTITUTE

COXETER LECTURE SERIES

MICHAEL JORDAN

University of California, Berkeley

APRIL 7 - 9, 2015 • THE FIELDS INSTITUTE, ROOM 230

APRIL 7, 3:30 P.M.

On Computational Thinking, Inferential Thinking and "Big Data"

The rapid growth in the size and scope of datasets in science and technology has created a need for novel foundational perspectives on data analysis that blend the inferential and computational sciences. That classical perspectives from these fields are not adequate to address emerging problems in "Big Data" is apparent from their sharply divergent nature at an elementary level---in computer science, the growth of the number of data points is a source of "complexity" that must be tamed via algorithms or hardware, whereas in statistics, the growth of the number of data points is a source of "simplicity" in that inferences are generally stronger and asymptotic results can be invoked. I present several research vignettes on topics at the computation/statistics interface, including the problem of trading off inference and privacy, the problem of inference under communication constraints and algorithmic weakening as a tool for trading off the speed and accuracy of inference.

APRIL 8, 3:30 P.M.

Lower Bounds at the Computational and Statistical Interface

One of the grand challenges of our era is the attempt to bring computational and statistical ideas together in a theoretically-grounded framework for scalable statistical inference. This is made challenging by the lack of a role for computational concepts such as "runtime" in core statistical theory and the lack of a role for statistical concepts such as "risk" in core computational theory. I discuss further attempts to build bridges between "computational thinking" and "inferential thinking," focusing on the theoretical study of lower bounds that embody computational and statistical constraints.

APRIL 9, 11:00 A.M.

Distributed Computing, the Bootstrap and Concurrency Control

Divide-and-conquer is a powerful paradigm in computer science, informing the design of algorithms and the design of distributed computing architectures. In the statistical setting, care must be taken, because naive use of divide-and-conquer can yield incorrect statistical inferences. Focusing on distributed computing architectures, I present a distributed version of the bootstrap and discuss the use of concurrency control to trade off the speed and accuracy of inference.

MICHAEL I. JORDAN is the Pehong Chen Distinguished Professor in the Department of Electrical Engineering and Computer Science and the Department of Statistics at the University of California, Berkeley. His research interests bridge the computational, statistical, cognitive and biological sciences, and have focused in recent years on Bayesian nonparametric analysis, probabilistic graphical models, spectral methods, kernel machines and applications to problems in distributed computing systems, natural language processing, signal processing and statistical genetics. Professor Jordan is a member of the National Academy of Sciences, a member of the National Academy of Engineering and a member of the American Academy of Arts and Sciences. He is a fellow of the American Association for the Advancement of Science. He has been named a Neyman Lecturer and a Medallion Lecturer by the Institute of Mathematical Statistics. He received the David E. Rumelhart Prize in 2015 and the ACM/AAAI Allen Newell Award in 2009.

For more information, please visit:
www.fields.utoronto.ca/programs/scientific/14-15/bigdata/CLS/

NSERC CRSNG

Ontario

THE FIELDS INSTITUTE FOR RESEARCH IN MATHEMATICAL SCIENCES
222 College Street, Second Floor, Toronto, Ontario, M5T 3J1 • www.fields.utoronto.ca • 416-348-9710