Faster Learning for Massive Datasets
Alexander Gray, Georgia Institute of Technology

I will describe new approaches for online learning and stochastic programming, which achieve both tighter theoretical bounds across the board and significant empirical gains over state-of-the-art approaches including stochastic gradient descent and mirror descent. I will then present a scheme for distributed online learning exhibiting first-of-a-kind theoretical and empirical gains. For nonlinear kernelized methods, kernel matrix multiplications and summations become a bottleneck. I will show fast algorithms which provably reduce computation times from quadratic to linear time, with corresponding empirical runtime results, demonstrated on over 10,000 cores.