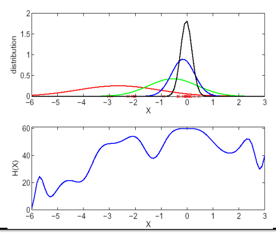
Simulation-Based Methods for Control and Optimization

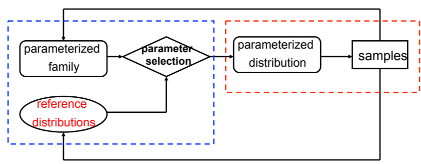
Michael Fu and Steve Marcus

ISR Faculty collaborators: Michael Ball, Rance Cleaveland, Jeffrey Herrmann, Dana Nau, Gary Rubloff

Model Reference Adaptive Search (MRAS) for Global Optimization

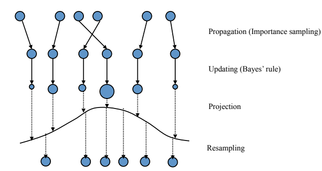
- Applicable to general global optimization problems
- Framework for design and analysis of algorithms
- Provably convergent and computationally efficient
- Probability distribution converges to distribution concentrated at the global optimum
- Applications: discrete optimization, continuous optimization, inventory control, buffer allocation, optimization problems in WDM networks, data mining

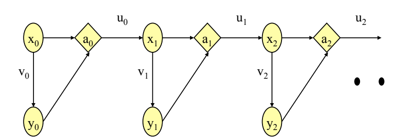




Simulation-Based Methods for Partially Observed Markov Decision Processes

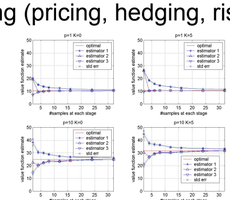
- Applicable to general continuous-state POMDPs
- Past work almost all for discrete-state POMDPs
- Provable error bounds and computationally efficient
- Incorporates new simulation-based Projection Particle Filter for state estimation
- Projects belief state onto exponential family
- Successful application to inventory control problem





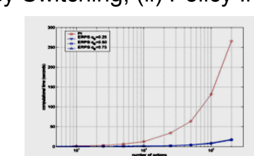
Adaptive Sampling Algorithm for Solving Markov Decision Processes

- Applicable to dynamic stochastic optimization problems that are difficult to model, or that have no analytical model but where simulations are available
- Goal: design & analysis of algorithms to estimate value function efficiently
- Use *adaptive sampling:* decide which action to sample next (from given state at given time)
    - How to decide: bandit model—tradeoff between exploration & exploitation: choose action that maximizes $\hat{Q}_i(x,a)+k\sqrt{\frac{\ln n_i}{n^x_{a,i}}}$
- Applications: manufacturing (preventive maintenance, capacity expansion in semiconductor fab), financial engineering (pricing, hedging, risk management)



Population Based Evolutionary Approaches for Solving Markov Decision Processes

- Large action space setting; alternative to policy improvement in policy iteration
- Goal: find optimal (or good) policies
- Approach: update *population* of policies as opposed to single policy
- Key ideas:
    - Avoid optimization over entire action space (parallelizable)
    - Monotonicity among elite policies
    - Convergence w.p. 1 to optimal value function
- Two methods of generating elite policy: (i) Policy Switching; (ii) Policy Improvement with Cost Swapping (PICS)



Support: NSF-ERC (ISR), NSF/SRC Operational Methods, NSF-DMI, NSF-CNS, NSF-CMMI, NSF-EECS, NSF-Expeditions in Computing, AFOSR