Research Overview

Michael Fu
University of Maryland

Workshop on
Decision Making in Adversarial Domains
May 23-25, 2005

Supported by NSF DMI-0323220, AFOSR FA95500410210

Projects

- Markov Decision Processes
  - with Steve Marcus
  - modeling and solution methodologies
  - simulation-based (sampling)
  - population-based (evolutionary, analytical)

covered today

- Simulation Methodology
  - perturbation analysis
  - computing budget allocation
  - importance sampling
- Computational Finance
  - pricing of American-style derivatives
  - credit risk (default)
- Others
  - fluid models: traffic network optimization, call centers

Personnel

- collaborative faculty
  - Steve Marcus (UM Electrical & Computer Engineering, ISR)
  - Dilip Madan (UM Business: Finance)
  - Jian-Qiang Hu (Boston University)
  - Chun-Hung Chen (George Mason University)
  - Xiaolan Xie (INRIA Metz, ENIM)
  - Dana Nau (UM Computer Science, ISR)

- 6 postdocs (over the years)

- 8 PhD students
  - 2 Operations Research/Management Science
  - 2 Electrical & Computer Engineering (control)
  - 3 Applied Mathematics
  - 1 Computer Science (working under Dana Nau)

Jiaqiao Hu

ROBERT H. SMITH SCHOOL OF BUSINESS

The Institute for Systems Research

A. JAMES CLARK SCHOOL OF ENGINEERING

Research goal for MDPs:
develop practically efficient computational methods

- evolutionary, population-based approaches
  - *large action space*
  - *complement state space reduction techniques (e.g., approx DP)*
  - *avoid optimization over the entire action space*
- simulation-based approaches
  - *large state space*
  - *transition probabilities not explicitly known or impractical to work with*
  - *no explicit mathematical model required*
- desired properties
  - *generic* and *robust* (for particular class of problems)
  - ***theoretical convergence*** guarantee
  - *good numerical results*

ROBERT H. SMITH SCHOOL OF BUSINESS
The Institute for Systems Research
A. JAMES CLARK SCHOOL OF ENGINEERING

Computational Approaches for Solving MDPs

(1) Population-based evolutionary algorithms

- Goal: find optimal **stationary** policy (infinite horizon)
- targeted at problems with **large action spaces** (possibly uncountable)
- departure from traditional approaches of policy iteration and value iteration

(2) Adaptive sampling approaches

- setting: transition probabilities not explicitly known, only samples; **finite horizon** problems
- targeted at problems with **large state spaces**
- limited sampling budget (e.g., simulation replications)
- **multi-armed bandit models** to decide which actions to sample

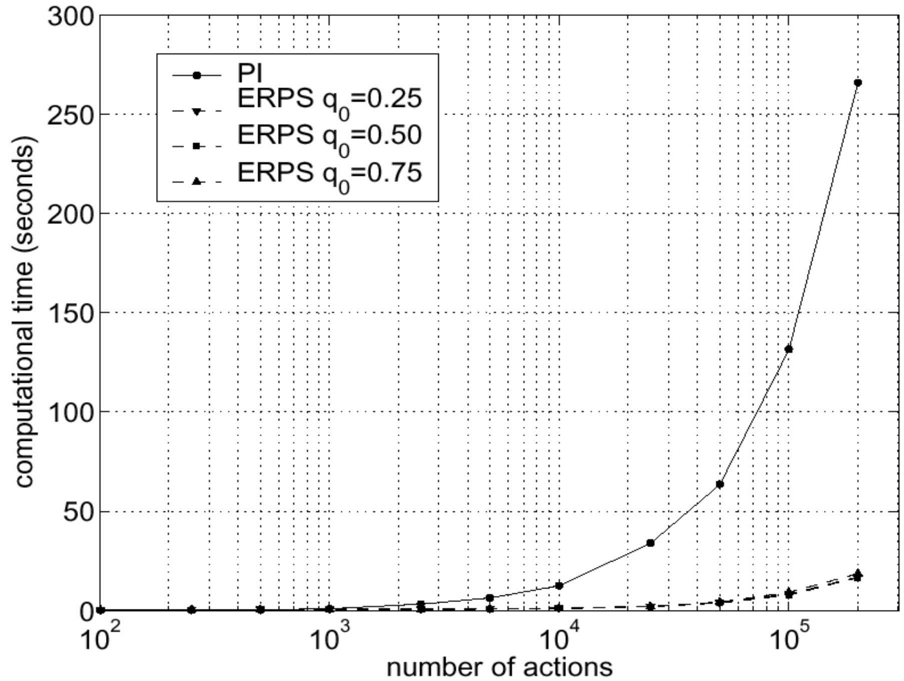
PhD student dissertation work of **Jiaqiao Hu**

Population-Based Approaches

- Chang, Lee, Fu, Marcus, "Evolutionary Policy Iteration for Solving Markov Decision Processes," *IEEE Transactions on Automatic Control* (submitted)
  1. Using **population** of policies (as opposed to iterating on a single policy) containing an "elite policy" based on "policy switching"
  2. Exploration and exploitation mechanisms provided
  3. Monotone property to guarantee finding (global) optimal policy, theoretical proof of convergence
  4. Conceptual and theoretical framework, no computational experiments

Population-Based Approaches (continued)

- Hu, Fu, Ramezani, Marcus, "An Evolutionary Random Search Algorithm for Solving Markov Decision Processes," *INFORMS Journal on Computing* (submitted)
  1. Combining modification of "policy switching" that works on a set of policies and local nearest neighbor search
  2. Promising experimental results, compared with standard policy iteration (PI) and previously proposed evolutionary algorithms (e.g., EPI)



Simulation-Based Approaches

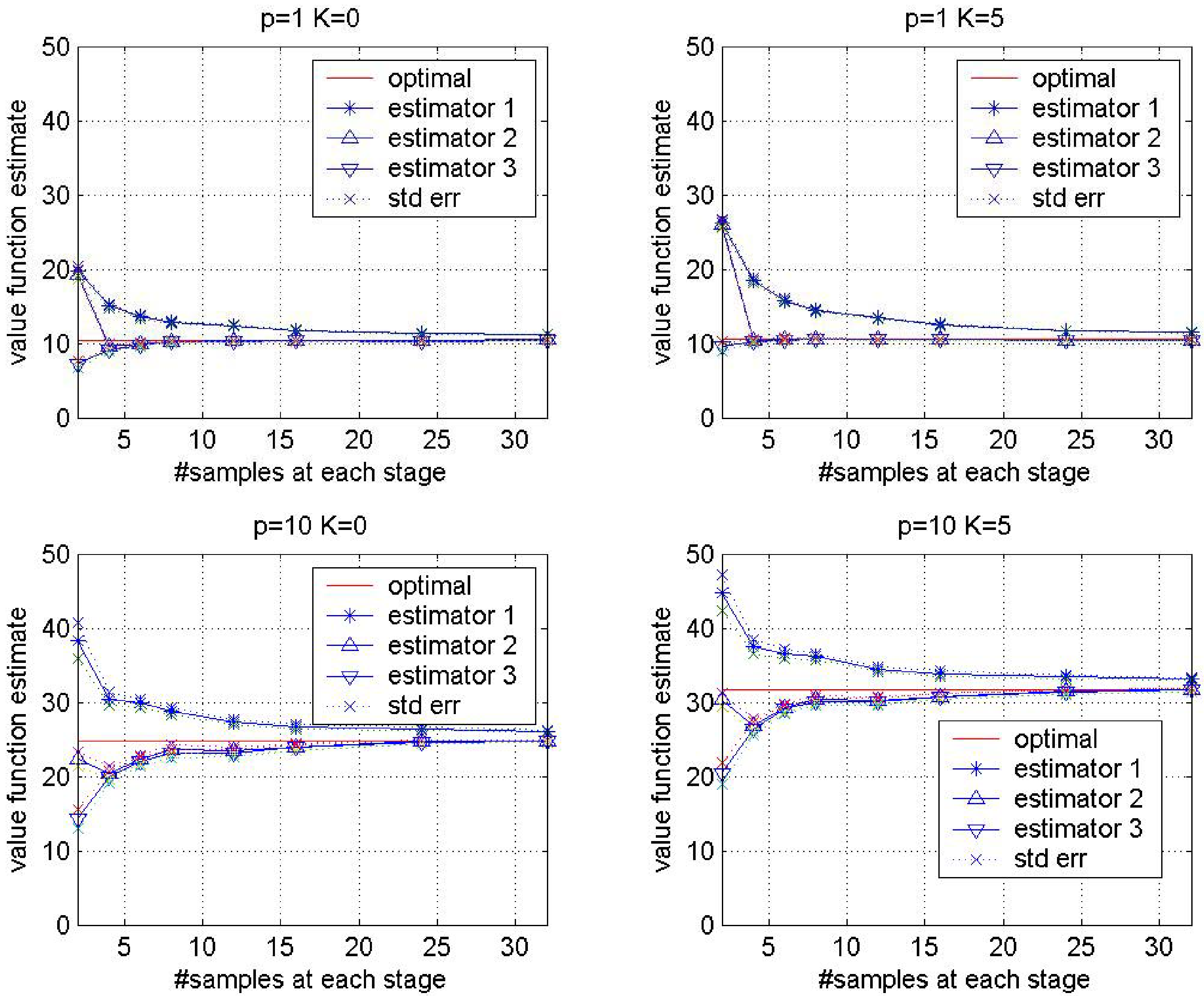
- Chang, Fu, Hu, Marcus, “An Adaptive Sampling Algorithm for Solving Markov Decision Processes,” *Operations Research*
- 3 different estimators
  - biased high
  - biased low
  - unknown (seems to be in between other two)
- bounded convergence rate
- worst-case complexity $O(N^H)$

  compare with backwards induction $O(H|A||X|^2)$
  - independent of size of state space X

    (action space A)
  - exponential in horizon length H

    (N = total number of simulation samples)

Numerical Results for an Inventory Control Example



New Project (funded by ISR Seed Grant)

- Planning Problems in AI
- Dana Nau and Ugur Kuter (from CS) & Jiaqiao Hu
- combine heuristic with adaptive sampling

- Action elimination
- Upper and Lower bounds
- EXPONENTIAL speed ups in some cases
- Example: UAV missions
- paper in preparation