An Adaptive Sampling Algorithm for Solving Markov Decision Processes

Abstract:

We present an adaptive sampling algorithm that approximates the optimal value of a finite-horizon Markov decision process (MDP) with finite action spaces.

The algorithm adaptively chooses which action to sample as the sampling process proceeds and generates an asymptotically unbiased estimator, whose bias is bounded by a quantity that converges to zero at rate (lnN)/N, where N is the total number of samples that are used per state sampled in each stage. To illustrate the algorithm, computational results are reported on simple examples from inventory control. Finally, we relate the algorithm to recent work on Monte Carlo tree search for playing Go.

Biography: Michael Fu is Ralph J. Tyser Professor of Management Science in the Robert H. Smith School of Business, University of Maryland at College Park, with a joint appointment in the Institute for Systems Research and affiliate faculty appointment in the Department of Electrical and Computer Engineering (in the A. James Clark School of Engineering). He received degrees in mathematics and EE/CS from MIT in 1985, and a Ph.D. in applied mathematics from Harvard University in 1989. His research interests include simulation optimization and applied probability, with applications in supply chain management and financial engineering. At Maryland, he received the Business School's Allen J. Krowe Award for Teaching Excellence in 1995, the Institute for Systems Research Outstanding Systems Engineering Faculty Award in 2002, and was named a University of Maryland Distinguished Scholar-Teacher for 2004-2005. He has published four books (all co-authored or co-edited):
Conditional Monte Carlo: Gradient Estimation and Optimization Applications, which received the INFORMS Simulation Society Outstanding Publication Award in 1998; Simulation-based Algorithms for Markov Decision Processes; Perspectives in Operations Research; and Advances in Mathematical Finance. He served as Stochastic Models and Simulation Department Editor of Management Science from 2006-2008, as Simulation Area Editor of Operations Research 2000-2005, and also on the editorial boards of Mathematics of Operations Research, INFORMS Journal on Computing, IIE Transactions, and Production and Operations Management. He is a Fellow of INFORMS and IEEE.