VOI-aware Monte Carlo Sampling in Trees

David Tolpin, Solomon Eyal Shimony
Department of Computer Science,
Ben-Gurion University of the Negev, Beer Sheva, Israel
{tolpin,shimony}@cs.bgu.ac.il

November 28, 2011

Abstract

Upper bounds for the VOI are provided for pure exploration in the Multi-armed Bandit Problem. Sampling policies based on the upper bounds are suggested. Empirical evaluation of the policies is provided on random problems as well on the Go game.
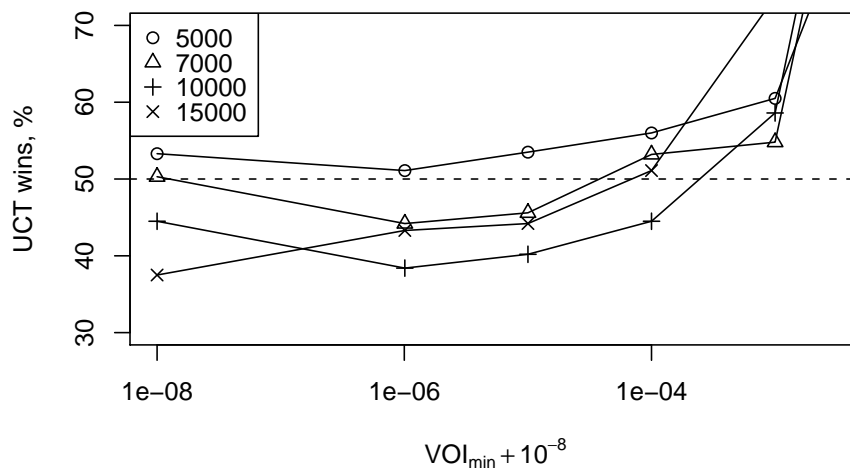
1 Empirical Evaluation

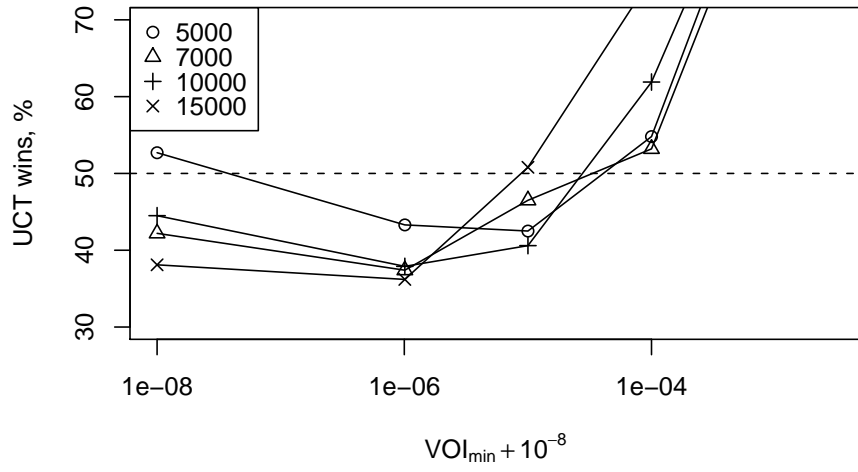Figure 1: Winning rate: UCT against VCT
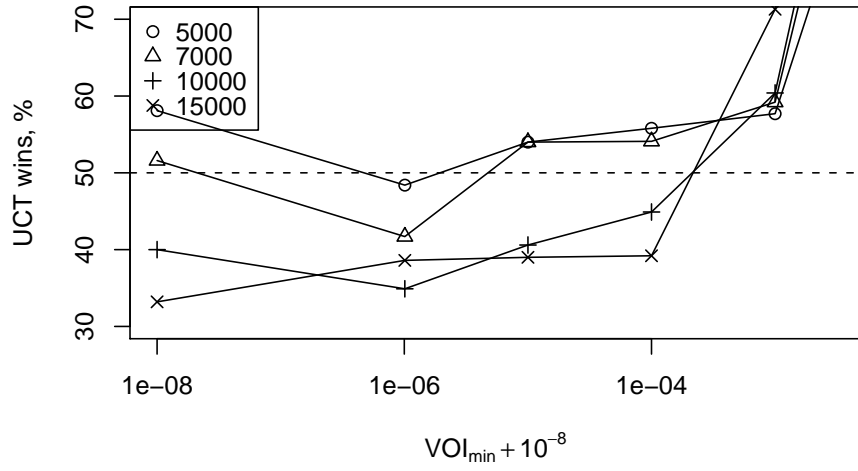
Figure 2: Winning rate: UCT against ECT
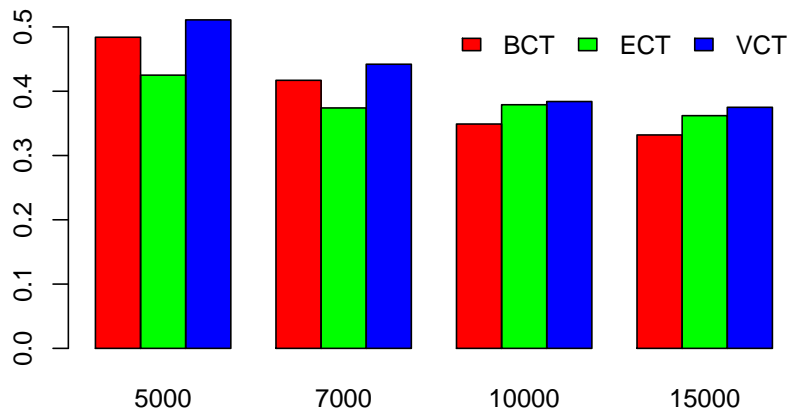

Figure 3: Winning rate: UCT against BCT

Figure 4: Best winning rate comparison

References