On Better Bound on VOI of Monte Carlo Sampling

David Tolpin, Solomon Eyal Shimony
{tolpin,shimony}@cs.bgu.ac.il

October 7, 2011

Consider the case of VOI of measurement the current best item (sample mean α). The probability that the true mean of the item μ_α is at most x is bounded by the Hoeffding's inequality as:

$$\Pr(\mu_\alpha \le x) \le \exp\left(-2n(\alpha - x)^2\right)$$

An upper bound on the VOI is thus the product of the probability that the true mean of the current best item is less than the sample mean β of the current second-best item times the maximum reward, β:

$$V_{\overline{x}=\alpha} \le \beta \exp\left(-2n(\alpha - \beta)^2\right)$$

The bound can be supposedly be improved by selecting a midpoint $0 < \gamma < \beta$ and computing the bound as the sum of two parts:

- $\beta - \gamma$ multiplied by the probability that $\mu_\alpha \le \beta$;
- β multiplied by the probability that $\mu_\alpha \le \gamma$.

.

$$V_{\overline{x}=\alpha} \le (\beta - \gamma) \exp\left(-2n(\alpha - \beta)^2\right) + \beta \exp\left(-2n(\alpha - \gamma)^2\right)$$

The minimum of V^* is achieved when $\frac{dV^*}{d\gamma} = 0$, that is, when γ is the root of the following equation:

$$4\beta n(\alpha - \gamma) = \exp\left(-2n\left(\frac{\alpha - \beta}{\alpha - \gamma}\right)^2\right)$$

If a root in the interval $0 \le \gamma \le \beta$ exists, then the number of samples is bounded as

$$n \le \frac{1}{4\beta(\alpha - \beta)}$$

by observing that the right-hand side is at most 1 (a negative power), and the left-hand side is at least $4\beta n(\alpha - \beta)$. So, the bound can supposedly be improved for smaller values of n. The improvement is more significant when the current best and second-best sample means are close.

The derivation for the other case (sampling an item that can be better than the current best) is obtained by substitution $1 - \overline{x}, 1 - \alpha, 1 - \gamma$ instead of α, β, γ:

$$V_{\overline{x}\ne\alpha} \le (\gamma - \alpha) \exp\left(-2n(\alpha - \overline{x})^2\right) + (1 - \alpha) \exp\left(-2n(\gamma - \overline{x})^2\right)$$

The anticipated influence of the improved VOI estimate would be that the selected item will be a less discovered one and further from the current best or second-best.

A closed-form solution for γ cannot be obtained, but given α, β, n, the value of γ can be efficiently computed. It should be determined empirically whether the improved estimate has justified influence on the performance of the algorithm.