

Reinforcement Learning Reading List

Mauricio Garcia Tec

Stephen G. Walker

Department of Statistics, University of Texas at Austin

November 21, 2018

Here's a list to Reinforcement Learning ideas and papers for personal research.

1 Key concepts

Multi-armed Bandits

It is the most elementary problem in RL, and the building block of many algorithms for more complex problems. We have random variables X_1, \dots, X_K representing the rewards of K possible actions, and we have to decide a sampling strategy so that we maximize the *total reward* over time. Basically, we are interested in $\max(X_1, \dots, X_K)$. We start with zero data from which to take a decision, and at each turn, we must select one of the X_i from which to sample (or one bandit arm—in analogy with casino machines). Of course, we could sample each arm 1,000 times and then choose whichever sample from these data has the highest mean; but the idea is to learn quickly and reach high values as soon as possible.

One of the key ideas in this body of literature is [Thompson Sampling][thompson] (1933). Which essentially boils down to Bayesian inference. We start with a uniform prior, and at each step we update our posterior belief of the arm selected by the algorithm with the observed value. We now have to choose according to which one we belief is the max of the variables, the usual approach is to simulate from each posterior, and the choose the one with the highest value.

Although the idea of Thompson sampling is very old, it remains an active area of research. For example, [Chappelle Li][chappelle-li] (2011) compares multiple strategies.

Another considerable part of the literature around the idea of minimizing the regret

$$\sum_{t=1}^T (\mu^* - \mu_{i(t)})$$

where μ^* is the mean of the X_i with highest mean, and $\mu_{i(t)}$ is the mean of the random variable chosen by the sampling strategy at time t .

****MonteCarlo Tree Search****

****Reinforcement Learning****

****Deep reinforcement learning****

—

Web

- [<https://spinningup.openai.com/>]() : It contains

—

Articles

****1****

****2****

****3****

****4****

****5****

****6****

****7****

****8****

****9****

****10****

****11****

****12****

****13****

****14****

References

- Chapelle Li An Empirical Evaluation of Thompson Sampling - Thompson, William R. "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples". Biometrika, 25(3-4):285-294, 1933.

[thompson]: [https://en.wikipedia.org/wiki/Thompson_sampling\[chappelle-li\]](https://en.wikipedia.org/wiki/Thompson_sampling[chappelle-li]) : <https://papers.nips.cc/paper/4321-an-empirical-evaluation-of-thompson-sampling.pdf>