Problem Set #3

Due: April 17, 2015

Dustin Tran, Xiaomin Wang, Rodrigo Gomes
{trandv,xiaominw, rgomes}@mit.edu

Natural scenarios in real life occur where one must sequentially make decisions under uncertainty. Quite often not only are transitions to certain states unknown but the true state of the agent as well. This is similar to that of a hidden Markov model (HMM), only in that one must make a sequence of actions instead of a single action. The classic scenario is a robot navigating a discrete environment using a GPS, and it takes action that lead to different states with various probabilities, but because of the GPS there is also inaccuracy in what its current state is.

One can formalize this, under Markovian assumptions, as a partially observable Markov decision process (POMDP). In this project we provide a robust software library for solving them with modular classes in order to allow for flexible extensions. More specifically, we encode a variety of basic tasks and solve them using a variant of Thompson sampling [1], which is a Bayesian approach following a Dirichlet-multinomial posterior over each state-action pair. The posterior probabilities are updated using the transition counts.

We follow the directory structure specified in the problem set, with two exceptions:

- `documentation/` does not exist. Instead, documentation is written in the `README.md` inside the current working directory. Any additional documentation not purely necessary for the problem set submission is in the Github wiki.
- `source/` is named `bayesrl/` in order to follow Python convention for installing modules.

References

[1] Malcolm Strens. A bayesian framework for reinforcement learning. In *Proceedings of the 17th International Conference on Machine Learning (ICML)*, 2000.