Thomas Kesselheim

November 6, 2018
Due: lecture on November 13, 2018

Algorithms and Uncertainty

Winter Semester 2018/19

Exercise Set 3

Exercise 1: (4 Points)
Let us consider a generalization of the version of Markov decision processes covered in the lecture. For every state $s \in \mathcal{S}$, only a subset of the actions $\mathcal{A}_s \subseteq \mathcal{A}$, $\mathcal{A}_s \neq \emptyset$, is available. Devise an algorithm that computes an optimal policy for a finite time horizon T, show its correctness, and give a bound on its running time.

Exercise 2: (1+5 Points)
We consider a stochastic decision problem similar to the one with the envelopes we solved in class. There are n boxes; box i contains a prize of 1 Euro with probability q_i and is empty otherwise. The game ends when we have found a non-empty box. That is, the final prize is either 0 Euros or 1 Euro. At each point in time, we can also decide to stop playing. We can open as many boxes as we like but opening box i costs c_i Euros.

(a) Model this problem as a Markov decision process. In particular, give the state and action sets as well as transition probabilities and rewards.

(b) Find an optimal policy.

Hint: It can be useful to consider the cases $n = 1$ and $n = 2$ first.

Exercise 3: (1+1+2+4 Points)
We generalize the optimal stopping problem with known distributions. Now, the policy is allowed to stop k times and collect the reward v_i each time it stops. (So, the reward is additive.)

(a) Model this problem as a Markov decision process.

(b) Give an example in which the optimal policy cannot be expressed by a sequence of thresholds $\tau_1, \ldots, \tau_n$.

(c) Find an appropriate generalization of Theorem 7.2.

(d) Generalize Theorem 7.4 and prove the generalization. As τ, use a threshold so that in expectation the sequence gets stopped $\frac{k}{2}$ times and make use of Markov's inequality.

Exercise 4: (2 Points)
Consider the cost-minimization variant of the optimal stopping problem. In step i, we can stop the sequence at cost c_i. We have to stop the sequence at some point and want to minimize the cost for doing so.
Show that there is **no** $\alpha < \infty$ such that for all distributions the optimal policy has cost at most $\alpha \mathbf{E}[\min_i c_i]$.