Thomas Kesselheim
Alexander Braun

May 13, 2020
Due: May 27, 2020 at noon

Algorithms and Uncertainty

Summer Term 2020

Exercise Set 4

Due to the public holiday (Christi Himmelfahrt) on Thursday, May 21, there will be no tutorial sessions next week. As a consequence, this sheet is only due in two weeks, but also covers content from the lecture on Monday, May 18 (exercises 4 + 5). We will discuss this sheet in the tutorials on May 28.

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: (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 $\alpha \mathbf{E}[\min_i c_i]$.

Hint: It sufficies to consider $n = 2$.

Exercise 4: (2 Points)
Consider the following distribution for the prize of box i: the prize v_i is equal to w_i with probability q_i and is 0 else. Compute the fair cap.

Exercise 5: (3 Points)

In order to generalize the Pandora's Box setup from the lecture, suppose we would like to match people $i \in [n]$ to boxes $j \in [m]$ (each person can take at most one prize home). We know that person i's value v_{ij} for the prize in box j is independently drawn from a distribution $\mathcal{D}_{ij}$, but it costs c_{ij} to inspect the exact value of the box v_{ij}. Consider A_{ij}, I_{ij}, σ_{ij}, κ_{ij} and b_{ij} to be the corresponding generalizations of the variables introduced in the lecture.

Show that for any policy π, the expected value is given by

$$V(\pi) = \sum_{i,j} \mathbf{E}\left[A_{ij}\kappa_{ij} - (I_{ij} - A_{ij})b_{ij}\right] \quad .$$