Exercise Set IX, Advanced Algorithms 2022

These exercises are for your own benefit. Feel free to collaborate and share your answers with other students. Solve as many problems as you can and ask for help if you get stuck for too long. Problems marked * are more difficult but also more fun :).

These problems are taken from various sources at EPFL and on the Internet, too numerous to cite individually.

1 **Alice, Bob and Charlie.** Suppose that Alice and Bob have two documents d_A and d_B respectively, and Charlie wants to learn about the difference between them. We represent each document by its word frequency vector as follows. We assume that words in d_A and d_B come from some dictionary of size n. Suppose further that the number of distinct words that occur in d_A or d_B but not in both documents is bounded by an integer $k \geq 1$.

Show that Alice and Bob can each send a $O(k \log^2 n)$-bit message to Charlie, from which Charlie can recover the words that occur in d_A or d_B but not in both with probability at least $9/10$. You may assume that Alice, Bob and Charlie have a source of shared random bits, and know the value of n.

2 Let $f : 2^N \to \mathbb{R}$ be a submodular function. Show that the following functions are also submodular:

- $g(S) = f(S \cup A)$ where A is a fixed set.
- $g(S) = f(S \cap A)$ where A is a fixed set.
- $g(S) = f(N \setminus S)$.

3 Consider a directed $G = (V, E)$ and define the set function $f : 2^V \to \mathbb{R}$ by

$$f(S) = |\{(u, v) \in E : u \in S, v \notin S\}| \qquad \text{for every } S \subseteq V\,.$$

That is, $f(S)$ equals the number of arcs that exits the set S.

3a Show that f is a (non-monotone) submodular function

3b Let S be a random subset of vertices obtained by including each vertex with probability $1/2$ independently of other vertices. Show that

$$\mathbb{E}[f(S)] = |E|/4 \geq \mathrm{OPT}/4\,,$$

where $\mathrm{OPT} = \max_{T \subseteq V} f(T)$.

Also give an example of a graph where $|E| = \mathrm{OPT}$ and thus it shows that the analysis is tight with respect to OPT.

3c (*) Consider any submodular function f that is

- non-negative: $f(T) \geq 0$ for all T.

Let S be a random subset of elements obtained by including each element with probability $1/2$ independently of other elements. Then

$$\mathbb{E}[f(S)] \geq \mathrm{OPT}/4\,,$$

where $\mathrm{OPT} = \max_T f(T)$.

This shows that the simple randomized algorithm actually gives a good approximation to any (even non-monotone) submodular function assuming it is non-negative.