CS-E3190 Principles of Algorithmic Techniques

07. Randomized Algorithms – Tutorial Exercise

1. **Karger's algorithm.** The following algorithm is a contraction algorithm for the minimum cut problem. The idea is to pick edges to contract (uniformly at random), until there are only two nodes left in the graph. The cut will then be given by the non contracted edges left at the end of the algorithm.

Algorithm 1: Karger's algorithm

```
input  : G = (V, E)
while |V| > 2 do
    Pick an edge e ∈ E uniformly at random.
    G ← G/e        (edge contraction step)
end
return The cut defined by G
```

We want to prove that the contraction algorithm outputs a minimum cut with probability at least $2/n(n-1)$. Let $C \subseteq E$ denote the edges of a specific minimum cut of size k.

(a) Prove that $|E| \geq nk/2$.

Solution. Observe that the minimum degree of G is k: otherwise the minimum degree vertex w would induce a smaller cut where one partitions would contain w and the other partition $G \setminus w$. Hence

$$|E| = 1/2 \cdot \sum_{v \in V} \mathrm{deg}(v) \geq 1/2 \cdot \sum_{v \in V} k = nk/2.$$

(b) Using part (a), upper bound the probability that the contraction algorithm picks an edge from C.

Solution. Because C has k edges, the probability of choosing one of them is $k/|E|$. From the previous question, we can deduce that the probability that the contraction algorithm picks an edge from C is

$$\frac{k}{|E|} \leq \frac{k}{nk/2} = \frac{2}{n}.$$

(c) Let P_n be the probability that the contraction algorithm on an n-vertex graph avoids C. Prove that $P_n \geq (1-2/n)P_{n-1}$ for $n \geq 3$.

Solution. For P_n, in the first iteration we avoid the edges of C with probability $1 - \frac{k}{|E|} \geq 1 - \frac{2}{n}$. After this, the contraction has $n-1$ vertices and the probability of avoiding the edges for the rest of the graph is P_{n-1}. Hence

$$P_n = (1 - k/|E|)P_{n-1} \geq (1 - n/2)P_{n-1}.$$

(d) Conclude the proof by bounding P_n.

Solution. We have $P_1 = P_2 = 1$ since the algorithm terminates when we have less than three vertices left. The previous recurrence can be expanded to

$$P_n \geq \prod_{i=0}^{n-3} \left(1 - \frac{2}{n-i}\right)$$

$$= \prod_{i=0}^{n-3} \frac{n-i-2}{n-i} = \frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdot \frac{n-4}{n-2} \cdots \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3}$$
$$= \frac{2 \cdot (n-2)!}{n!} = \frac{2}{n(n-1)},$$

completing the proof.

2. **Balls and bins**[1]**.** The following problem models a simple distributed system wherein agents contend for resources but "back off" in the face of contention. Balls represent tasks and bins represent computers. The system evolves over rounds. Every round, balls are thrown independently and uniformly at random into n bins. Any ball that lands in a bin by itself is served and removed from consideration. The remaining balls are thrown again in the next round. We begin with n balls in the first round, and we finish when every ball is served.

   (a) Suppose there are b balls at the start of a round. Show that the expected number of balls at the start of the next round is at most $\frac{b(b-1)}{n}$.

   Solution. For each ball i let X_i be the indicator random variable for the event that i lands in a bin by itself. The total number of configurations of balls in bins is n^b. The number of configurations for which $X_i = 1$ is $n \cdot (n-1)^{b-1}$. We have $\mathrm{E}[X_i] = \Pr[X_i = 1] = (\frac{n-1}{n})^{b-1} = (1 - \frac{1}{n})^{b-1}$

   The number of balls at the start of the next round is $b - \sum_{i=1}^{b} X_i$, and therefore, the expected number of balls in the next round is:

$$b - \sum_{i=1}^{b} \left(1 - \frac{1}{n}\right)^{b-1} \leq b - b\left(1 - \frac{b-1}{n}\right) = \frac{b(b-1)}{n}$$

   Where the second inequality is due to the fact that $(1-x)^m \geq 1-mx$ for all $0 \leq x \leq 1$.

   (b) Suppose that every round the number of balls served was exactly the expected number of balls to be served. Show that all the balls would be served in $O(\log\log n)$ rounds.

   Hint: If x_j is the expected number of balls left after j rounds, notice that $x_{j+1} \leq x_j^2/n$.

   Solution.

   By part (a) we have $x_{j+1} \leq \frac{x_j(x_j-1)}{n} \leq \frac{x_j^2}{n}$. Because we have $x_1 = n$, this inequality gives us $x_2 \leq n$ which is not very useful as we don't make any progress.

   The trick is to observe that if $x_1 \leq n/2$, then the inequality gives us $x_2 \leq n/2^2$, $x_3 \leq n/2^4$, $x_4 \leq n/2^8$, and generally $x_{j+1} \leq n/2^{2^j}$. Now for $j = \log\log n$, we have $x_{j+1} \leq 1$, and we are done after $j + 2 = O(\log\log n)$ rounds.

   Now, when $x_1 = n$, how do we get to a point where half of the balls are processed? From the previous part we have

$$x_2 = n - n\left(1 - \frac{1}{n}\right)^{n-1} \leq n - n\left(1 - \frac{1}{n}\right)^{n} \leq n - \frac{n}{\sqrt{e}}$$

   where for the last inequality we use the fact that $1 - x \geq e^{x/2}$ for all $0 \leq x \leq 1$.

   Therefore, we have $x_{j+1} \leq x_j(1 - 1/\sqrt{e}) \leq x_1(1 - 1/\sqrt{e})^j$, and if we put $j = \log_{1/(1-1/\sqrt{e})} 2$, we get $x_{j+1} \leq x_1/2$.

   In other words, after a constant number of rounds, we expect at least half of the balls to be processed and then we can switch to the analysis that assumes $x_1 \leq n/2$.

[1] Exercise 5.12 Mitzenmacher Upfal Book