Notes 1 - Solutions

PROBLEM 1:

Suppose $k = n/2$. Show that the number of possible subsets of size k is at least $2^{\Omega(n)}$.

Divide V at random into two disjoint subsets $X, Y \subset V$ of size $\frac{n}{2}$.

Label the elements of these sets $\{x_1, ...x_{\frac{n}{2}}\}$ and $\{y_1, ...y_{\frac{n}{2}}\}$ respectively so that each element $x_i \in X$ has a natural pair $y_i \in Y$ for $i = \{1, ..., \frac{n}{2}\}$.

In order to form a subset $S \subseteq V$ of size $\frac{n}{2}$ choose exactly one of x_i or y_i for all $i = \{1, ..., \frac{n}{2}\}$.

There are therefore 2 choices for each element of S, implying that there are $2^{\frac{n}{2}}$ possible subsets of V of size $\frac{n}{2}$ formed in this way.

This implies that there are at least $2^{\Omega(n)}$ possible subsets of size $\frac{n}{2}$

PROBLEM 2:

What is the range of possible values for the approximation factor of an approximation algorithm for the influence maximization problem

This is a maximization problem, therefore, approximation factor $C \leq 1$. Additionally, in case of an algorithm that outputs k unique nodes, $k \leq f(x) \leq OPT$ and $\frac{k}{OPT} \leq C \leq 1$.

PROBLEM 3:

Suppose we want to find shortest paths. Is this is a maximization or minimization problem? What is the range of possible values for the approximation factor of an approximation algorithm for this problem

This is a minimization problem, we wish to find a set x of edges such that the source and destination nodes are connected by these edges, but the size of this set is minimal. The approximation factor C will satisfy $C \geq 1$.

If we want to be more technical, we can observe that a "path" cannot have repeated edges. Thus in an unweighted graph, the output of the algorithm has at most E edges, and C is bounded by $1 \leq C \leq \frac{E}{OPT}$.

PROBLEM 4:

Suppose a tortoise is at distance n meters from its destination. It is getting tired with time, and in each hour, it covers $\frac{1}{2}$ of the remaining distance to the destination. How long does it take the tortoise to get to the destination

At time step 0 the tortoise has a distance n left to travel. At the next time step a distance of $\frac{n}{2}$ left to travel, and so on.

Therefore the sequence representing the distance remaining is given by $\{n, \frac{n}{2}, \frac{n}{4}, ...\frac{n}{2^t}\}$ where t represents the time step in hours. At any time step this value will be a fraction with a value greater than 0, therefore the time to the destination is infinite.

PROBLEM 5:

How long does it take it to get to within 1 meter of the destination?

This occurs when $\frac{n}{2^t} < 1$ therefore $t > \log_2(n)$.

PROBLEM 6:

Now suppose instead the tortoise covers $\frac{1}{k}$ of the remaining distance to the destination in each hour. How long does it take it to get to within 1 meter?

In this case the sequence of distances remaining can be represented by:

$$\{n, \frac{(k-1)n}{k}, ..., \frac{(k-1)^t n}{k^t}\} \tag{1}$$

Therefore the distance is within 1 meter when $\frac{(k-1)^t n}{k^t} < 1$.

Solving for t we obtain $t > \frac{\log_k(n)}{1-\log_k(k-1)}$

PROBLEM 7:

If the tortoise covers $\frac{1}{k}$ of the remaining distance to the destination in each hour, what fraction of the distance remains to be covered after 1 hr? What fraction remains to be covered after k hours?

After 1 hour the tortoise has $\frac{(k-1)n}{k}$ meters left to cover. Therefore the fraction of distance left to cover is $\frac{\frac{(k-1)n}{k}}{n} = \frac{(k-1)}{k}$

After k hours the tortoise has $\frac{(k-1)^k n}{k^k}$ meters left to cover, so the fraction of the distance left to cover is $\left(1 - \frac{1}{k}\right)^k \leq \frac{1}{e}$.

PROBLEM 8:

In the independent activation model, suppose node u has neighbors $v_1, v_2, ...$ and the corresponding edges have associated probabilities $p_1, p_2, ...$. Suppose the probability of u being activated is $p(u)$. If all neighbors $v_1, v_2, ...$ are active, can we say that $p(u) \leq p_1 + p_2 + ...$? What is the exact probability of u becoming active?

In order for a node to become active, it must be activated by at least one of its neighbors. Therefore it is activated if it is activated by neighbour v_1 or neighbour v_2 and so on... We are in the independent activation model therefore $P(u)$ is given by (by union bound):

$$\begin{aligned} P(\text{activated by } v_1 \text{ OR activated by } v_2 \text{ OR } ...) &= P(\text{activated by } v_1) + P(\text{activated by } v_2) + ... \\ &= p_1 + p_2 + ... \end{aligned}$$

To compute the exact probability, we observe that the probability that u is not activated by v_i is $1 - p_i$. Thus the probability that u is not activate by any of the neighbors is $(1-p_1)(1-p_2)\ldots$. Thus the probability that u is activated by at least one neighbor is: $1 - \Pi_i(1-p_i)$.

PROBLEM 9:

Suppose for a node x in a network, an edge to each other $n-1$ nodes exists with probability $\frac{\ln(n)}{n-1}$. Show that the probability that x has no edge is $\leq \frac{1}{n}$ [hint: use $(1-\frac{1}{x})^x$ with $x = \frac{\ln(n)}{n-1}$.]

For each of the $n-1$ other nodes in the network, the probability that node x is not connected to it is $1 - \frac{\ln(n)}{n-1}$.

Assuming that the probability of node x being connected to any other nodes is independent of the other nodes it is connected to, then:

$$P(x \text{ is connected to no other node}) = \left(1 - \frac{\ln(n)}{n-1}\right)^{n-1}$$

Let $z = \frac{n-1}{\ln(n)}$, then:

$$\begin{aligned} P(x \text{ is connected to no other node}) &= \left[\left(1 - \frac{1}{z}\right)^z\right]^{\frac{1}{\ln(n)}} \\ &\leq \left(\frac{1}{e}\right)^{\frac{1}{\ln(n)}} \\ &= \frac{1}{e^{\ln(n)}} \\ &= \frac{1}{n} \end{aligned}$$