School of Computer Science and Technology
University of Science and Technology of China

Exercise Sheet 5 for
Design and Analysis of Algorithms
Autumn 2022
Solution

Exercise 1 **(30 points, graded by Yudong Zhang)**
Recall the (Minimum) Vertex Cover and (Maximum) Independent Set problems. For the Vertex Cover problem, give a graph $G = (V, E)$, the goal is to find a smallest subset $S \subseteq V$ such that for any edge $(u, v) \in E$, either $u \in S$ or $v \in S$. For the Independent Set problem, given a graph $G = (V, E)$, the goal is to find a largest subset $I \subseteq V$ such that there is no edge between any two vertices in I.
A simple observation is: if S^* is the *minimum* vertex cover in G, then $V \setminus S^*$ is *maximum* independent set in G.
Suppose you have given a 2-approximation algorithm for Vertex Cover. Consider the following approximation algorithm for Independent Set problem: Given a graph $G = (V, E)$, use the 2-approximation algorithm for Vertex Cover on G to get a vertex cover S and output $V \setminus S$.

- Give an upper bound on the approximation ratio of the above algorithm for the Independent Set problem. (The smaller your bound is, the better.)

  Solution. The 2-approximation algorithm for Vertex Cover outputs a vertex cover S which satisfies that $|S| \leq 2|S^*|$. Thus, $|I| = |V \setminus S| = |V| - |S| \geq |V| - 2|S^*|$. The approximation ratio of the above algorithm for the Independent Set problem is $\frac{|I|}{|I^*|} = \frac{|V \setminus S|}{|V \setminus S^*|} = \frac{|V|-2|S^*|}{|V|-|S^*|}$. It is easy to construct an example in which $|S^*| = \frac{|V|}{2}$, then we obtain that the above approximation ratio is at most 0. This also implies that a 2-approximation algorithm for Vertex Cover does not give any useful information for Independent Set problem even though $|S^*| + |I^*| = |V|$.

Exercise 2 **(30 points, graded by Yinhao Dong)**
Consider the LCR algorithm for the leader election problem on the ring network.

- Give a UID assignment for which $\Omega(n^2)$ messages are sent. **(15 points)**

  Solution. Assign the UIDs clockwise in *decreasing* order (e.g., $n, n-1, \ldots, 1$) so that the message containing UID i must be passed i times. Thus, the total number of messages sent is $n + (n-1) + \cdots + 1 = \sum_{i=1}^{n} i = \frac{n(n+1)}{2} = \Omega(n^2)$.

- Give a UID assignment for which only $O(n)$ messages are sent. **(15 points)**

  Solution. Assign the UIDs clockwise in *increasing* order (e.g., $1, 2, \ldots, n$) so that each message (except the message containing $u_{\max}$) only goes once. There are $n-1$ of these, while the message containing $u_{\max}$ requires n passes. Thus, the total number of messages sent is $n + (n-1) = 2n - 1 = O(n)$.

Exercise 3 **(40 points, graded by Di Wu)**
Consider a regularized variant of Luby's MIS algorithm as follows: The algorithm consists of $\log \Delta + 1$ phases, each made of $100 \cdot \log n$ consecutive rounds. Here Δ denotes the maximum degree in the graph. Let S be a set initialized to be empty. In each round of the i-th phase, each remaining node is marked with probability $\frac{2^i}{10\Delta}$. Different nodes are marked independently. Then marked nodes who do not have any marked neighbor are added to S, and removed from the graph along with their neighbors. If at any time, a node v becomes isolated and none of its neighbors remain, then v is also added to S and is removed from the graph.

- Argue that the set S at the end of the algorithm is an independent set. **(10 points)**

Proof. By contradiction. Suppose that two neighboring nodes u and v are added to S. Without loss of generality, we suppose that u was added no later than v, and let t be the round in which u was added to S. On the one hand, node v could not have been added in the same round t, because then u and v would be two neighboring marked nodes, and thus neither would be added. On the other hand, node v could not have been added in any round strictly after t, because at the end of round t, node u gets removed from the graph along with all of its neighbors, including v. We obtain a contradiction. Clearly, the condition that isolated nodes are added to S cannot change the fact that S is indeed an independent set. $\square$

- Prove that with high probability (i.e., with probability at least $1 - \frac{1}{n^C}$, for some constant $C > 1$), by the end of the i-th phase, in the remaining graph each node has degree at most $\frac{\Delta}{2^i}$. **(20 points)**

  Proof. By induction. The base case $i = 0$ is trivial. Consider the time at the beginning of the i-th phase, and suppose that each remaining node has degree at most $\frac{\Delta}{2^{i-1}}$. Consider an arbitrary node v and suppose that v has at least $\frac{\Delta}{2^i}$ remaining neighbors, at the beginning of this phase.

  We want to bound the probability that v remains with at least $\frac{\Delta}{2^i}$ neighbors by the end of the i-th phase. Per round of this phase, either at most $\left(\frac{\Delta}{2^i} - 1\right)$ neighbors of v remain (or v gets removed), in which case we are done, or at least $\frac{\Delta}{2^i}$ neighbors of v remain. In the latter case, there is a constant probability that, in the following round, (1) a neighbor u of v gets marked and (2) no neighbor of u and no neighbor of v (except u) gets marked. The probability that (1) holds is $\frac{2^i}{10\Delta}$. The probability that (2) holds is at least $\left(1 - \frac{2^i}{10\Delta}\right)^{2\cdot\frac{\Delta}{2^{i-1}}} = \left(1 - \frac{2^i}{10\Delta}\right)^{\frac{10\Delta}{2^i}\cdot\frac{2}{5}} > \left[\left(1 - \frac{2^i}{10\Delta}\right)^{\frac{10\Delta}{2^i}-1}\right]^{\frac{2}{5}} \geq e^{-\frac{2}{5}} > 4^{-\frac{2}{5}} > \frac{1}{2}$ since the fact that $\left(1 - \frac{1}{x}\right)^{x-1} \geq \frac{1}{e}$ for $x \geq 1$. Therefore, the probability that a specific neighbor u of v satisfies these two properties is at least $\frac{2^i}{10\Delta} \cdot \frac{1}{2}$, which implies that at least one neighbor of v has these properties is at least $\frac{\Delta}{2^i} \cdot \frac{2^i}{10\Delta} \cdot \frac{1}{2} = \frac{1}{20}$. (Note that it cannot be the case that two different neighbors of v both satisfy these properties; hence, we can simply add up the probabilities for different neighbors of v without causing over counting.) We conclude that in each round of the i-th phase, node v with degree at least $\frac{\Delta}{2^i}$ gets removed (due to having a neighbor added to S) with probability at least $\frac{1}{20}$.

  Considering the $100 \log n$ rounds of the phase, the probability of v remaining with degree at least $\frac{\Delta}{2^i}$ by the end of the i-th phase is at most $\left(1 - \frac{1}{20}\right)^{100\log n} = \left(1 - \frac{1}{20}\right)^{20\cdot 5\log n} \leq e^{-5\log n} < 2^{-5\log n} = \frac{1}{n^5}$ since the fact that $\left(1 - \frac{1}{x}\right)^x \leq \frac{1}{e}$ for $x \geq 1$. A union bound over all such nodes v shows that with probability at least $1 - \frac{1}{n^5} \cdot n \geq 1 - \frac{1}{n^4}$, no such node with degree at least $\frac{\Delta}{2^i}$ remains by the end of the i-th phase (assuming that each remaining node at the beginning of the i-th phase has degree at most $\frac{\Delta}{2^{i-1}}$). Since we have $\log\Delta + 1$ phases, and in each phase the probability that something goes wrong (and that in all previous phases nothing went wrong) is at most $\frac{1}{n^4}$, the probability that something goes wrong in *at least one phase* is at most $\frac{\log\Delta+1}{n^4} \leq \frac{1}{n^3}$. Hence, the statement we want to prove indeed holds with high probability (i.e., with probability at least $1 - \frac{1}{n^3}$). $\square$

  Note: Another approach (`https://disco.ethz.ch/courses/fs18/podc/exercises/solution5.pdf`) is also OK.

- Conclude that the set S at the end of the algorithm is a maximal independent set, with high probability. **(10 points)**

  Proof. By what we have proved above, by the end of phase $\log\Delta + 1$, in the remaining graph each node has degree at most $\frac{\Delta}{2^{\log\Delta+1}} = \frac{\Delta}{2\Delta} = \frac{1}{2}$, with high probability. This means the degree is actually 0. Once a node reaches degree 0, it gets added to the S. If the node v was removed anytime before that, it must have been that v was added to S or a neighbor of v was added to S. Therefore, S at the end of the algorithm is a maximal independent set, with high probability (i.e., with probability at least $1 - \frac{1}{n^3}$). $\square$
