COSC-373: HOMEWORK 4

LEE JIAEN, HYERY YOO, AND ALEXANDER LEE

Question 1. Each node v executes as follows.

Each round:

(1) Pick a uniformly random color $c(v)$ from the range $1, 2, \ldots, 2\deg(v)$
(2) Send message $c(v)$ to all neighbors
(3) If $c(v) \neq c(u)$ for all neighbors u, output $c(v)$ and halt

(Correctness) At each round a node v picks a color $c(v)$ uniformly at random from the range $1, 2, \ldots, 2\deg(v)$. Then, node v sends its chosen color $c(v)$ to all its neighbors. Similarly, node v would also receive colors from its active neighbors. If one of v's neighbors has already halted, v uses the last color it received from the neighbor. For each neighbor u of node v, if $c(v) \neq c(u)$, the condition that neighboring vertices are assigned different colors is satisfied, so node v outputs $c(v)$ and halts. Therefore, this protocol produces a proper 2Δ coloring of G.

(Congestion) Since the protocol only requires nodes to send colors from the range $1, 2, \ldots, 2\Delta$ to neighboring nodes, only messages of size $O(1)$ bits are being sent. Thus, the protocol is in the CONGEST model.

(Runtime) To show that the protocol produces a proper 2Δ coloring of G in $O(\log n)$ rounds, first define the random variable X_v for whether a node v halts in a given round. Specifically, X_v is 1 if node v halts in the given round, and 0 otherwise. For a given round, since node v has $\deg(v)$ neighbors, its neighbors can choose at most $\deg(v)$ distinct colors from each other. We also know that node v chooses uniformly at random from $2\deg(v)$ colors. As such, the probability that node v chooses the same color as one of its neighbors, is at most $\frac{\deg(v)}{2\deg(v)} = \frac{1}{2}$. It follows that the probability $\mathbb{P}(X_v)$ that node v halts in a given round (i.e., choosing a different color from its neighbors) is

$$\begin{aligned}\mathbb{P}(X_v) &\geq 1 - \frac{1}{2} \\ &= \frac{1}{2}.\end{aligned}$$

Hence, the expected value $\mathbb{E}(X_v)$ of node v halting in a given round is

$$\begin{aligned}\mathbb{E}(X_v) &\geq 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} \\ &= \frac{1}{2}.\end{aligned}$$

Next, define the random variable X for the number of nodes that halt in a given round. Given the set of nodes V for graph G, we have that

$$X = \sum_{v \in V} X_v \ .$$

Taking the expected value on both sides, we get

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}\left(\sum_{v \in V} X_v\right) \\ &= \sum_{v \in V} \mathbb{E}(X_v) \\ &\geq \sum_{v \in V} \frac{1}{2} \\ &= \frac{1}{2} n\end{aligned}$$

Thus, in expectation, at least $\frac{1}{2}$ of all nodes halt in each round. If $n = n_0, n_1, \ldots, n_k$ are the number of active nodes after rounds $0, 1, \ldots, k$, we have that

$$\mathbb{E}(n_k) \leq \frac{1}{2^k} \cdot n.$$

Taking $k = 4 \log n$, we get

$$\begin{aligned}\mathbb{E}(n_k) &\leq \frac{1}{2^{4 \log n}} \cdot n \\ &= \frac{1}{n^4} \cdot n \\ &= \frac{1}{n^3}.\end{aligned}$$

By Markov's Inequality, it follows that $\mathbb{P}(n_k \geq 1) \leq \frac{1}{n^3}$. That is, after $4 \log n$ rounds, the probability that there are one or more active nodes is at most $\frac{1}{n^3}$. Therefore, the protocol produces a proper 2Δ coloring of G in $O(\log n)$ rounds with high probability.

Question 2. For $F = \{(A, A^c) \mid A \subseteq \{1, 2, \ldots, N\}\}$ to be a fooling set for function DISJ, we must prove the following conditions:

(1) For all i, j, we have DISJ$(A_i, A_i^c) =$ DISJ(A_j, A_j^c); and
(2) For all $i \neq j$, we have DISJ$(A_i, A_i^c) \neq$ DISJ(A_i, A_j^c) or DISJ$(A_i, A_i^c) \neq$ DISJ(A_j, A_i^c).

To prove condition (1), let $(A_i, A_i^c) \in F$. Since A_i^c is the complement of A_i, it follows that A_i and A_i^c are disjoint because $A_i \cap A_i^c = \emptyset$. Thus, DISJ$(A_i, A_i^c) = 1$. Notice that (A_i, A_i^c) is any arbitrary element in F. Therefore, for all i, j where $(A_i, A_i^c), (A_j, A_j^c) \in F$, we have that DISJ$(A_i, A_i^c) = 1 =$ DISJ(A_j, A_j^c).

To prove condition (2), let $(A_i, A_i^c), (A_j, A_j^c) \in F$, where $i \neq j$. Three cases are possible.

Case 1: $A_j \subset A_i$ (and $A_j \not\subset A_i^c$). This implies that $A_i \cap A_j^c \neq \emptyset$ and $A_j \cap A_i^c = \emptyset$. Thus, we have that DISJ$(A_i, A_j^c) = 0$ and DISJ$(A_j, A_i^c) = 1$. Since DISJ$(A_i, A_i^c) = 1 \neq 0 =$ DISJ(A_i, A_j^c), condition (2) is satisfied for this case.

Case 2: $A_j \subset A_i^c$ (and $A_j \not\subset A_i$). This implies that $A_i \cap A_j^c \neq \emptyset$ and $A_j \cap A_i^c \neq \emptyset$. Hence, we have that DISJ$(A_i, A_j^c) = 0$ and DISJ$(A_j, A_i^c) = 0$. Since DISJ$(A_i, A_i^c) = 1 \neq 0 =$ DISJ(A_i, A_j^c), condition (2) is also satisfied for this case.

Case 3: $A_j \not\subset A_i$ and $A_j \not\subset A_i^c$. This implies that $A_i \cap A_j^c \neq \emptyset$ and $A_j \cap A_i^c \neq \emptyset$. It follows that DISJ$(A_i, A_j^c) = 0$ and DISJ$(A_j, A_i^c = 0)$. Since DISJ$(A_i, A_i^c) = 1 \neq 0 =$ DISJ(A_i, A_j^c), condition (2) is satisfied for this case, and the proof is complete.

Question 3. The table below shows the result of GT for values x and y between 0 and $2^N - 1$, where $N = 3$. The x values are on the columns and y values are on the rows.

	0	1	2	3	4	5	6	7
0	0	1	1	1	1	1	1	1
1	0	0	1	1	1	1	1	1
2	0	0	0	1	1	1	1	1
3	0	0	0	0	1	1	1	1
4	0	0	0	0	0	1	1	1
5	0	0	0	0	0	0	1	1
6	0	0	0	0	0	0	0	1
7	0	0	0	0	0	0	0	0

In general, given an arbitrary N and $0 \le x, y \le 2^N - 1$, $\mathrm{GT}(x, y)$ is 0 on the diagonal of the table since $x = y$, 0 below the diagonal since $x < y$, and 1 above the diagonal since $x > y$. It follows that a fooling set for GT is the diagonal of the table where $x = y$, i.e., $F = \{(0,0), (1,1), (2,2), \ldots, (2^N - 1, 2^N - 1)\}$. F is a fooling set for GT because

(1) for all i, j, $\mathrm{GT}(x_i, y_i) = \mathrm{GT}(x_j, y_j) = 0$; and
(2) for all $i > j$, $\mathrm{GT}(x_i, y_i) = 0 \neq 1 = \mathrm{GT}(x_i, y_j)$, and for all $i < j$, $\mathrm{GT}(x_i, y_i) = 0 \neq 1 = \mathrm{GT}(x_j, y_i)$. Combining the two cases (when $i > j$ and $i < j$), we have that for all $i \neq j$, $\mathrm{GT}(x_i, y_i) \neq \mathrm{GT}(x_i, y_j)$ or $\mathrm{GT}(x_i, y_i) \neq \mathrm{GT}(x_j, y_i)$.

Observe that $|F| = 2^N$. The fooling set method states that for a function f and a fooling set F of size l, $D(f) \ge \log(l)$. Applying this to GT, $D(GT) \ge \log(|F|) = \log(2^N) = N$. Therefore, the deterministic communication complexity of GT is at least N.