CSE638 & AMS641: Advanced Algorithms **Lecturer:** Prof. Rezaul A. Chowdhury
Submitted by: Yoni Fogel, Arun Rathakrishnan and Amitav Paul **Date:** April 20, 2013

Homework # 2

1 Parallel BFS with Cilk's Work Stealing.

1.a

We replace the **for** loop over an input queue with a manually implemented **parallel for** (divide-and-conquer) and used the worker's id to choose the output queue.

1.b

Input File	RT_{SBFS}	$RT_{PBFS^{(f)}}$	$RT_{PBFS^{(g)}}$	$SF_{PBFS^{(f)}}$	$SF_{PBFS^{(g)}}$	T_{HW2}
cage14	1786	313	315.18	5.70	5.67	55.65
cage15	1231	303.904	303.75	4.05	4.05	212.18
freescale	1835	223	253.00	8.22	7.25	218.56
Wikipedia	709	705.47	703	1.00	1.00	139.19
kkt-power	410	21.5	20.79	19.07	19.72	10.06
RMAT100M	3253	673.57	647.00	4.83	5.03	501.59
RMAT1B	-	2544.57	-	-	-	2400

Table 1: RT = Running Time, SBFS = Serial BFS, $PBFS^{(f)}$ = Parallel BFS for task 1(f), $PBFS^{(g)}$ = Parallel BFS for task 1(g) and HW2 = Homework 2. Times are in seconds.

2 Parallel Connected Components.

2.a

For an arbitrary graph that satisfies the required property let us assume for the sake of contradiction, the largest possible Q is such that $|Q| < \frac{|V|}{2}$. $G' = (V, Q')$ is cycle free. Adding more edges will lead to the formation of a cycle in G'.

Consider all the potential edges that can be added to graph $E' = \{(v, N[v]), v \in V\}$. $|E'| = |V|$, since each vertex u, has a unique $N[u]$. Clearly $|Q| < \frac{|E'|}{2}$.

If there are k edges in a graph, at most $2 \times k$ vertices can have an edge incident on them. So G' has less than $|V|$ vertices that have an edge incident on them. There is atleast one non-isloated vertex v, that has no edge incident on it in G'.

The partition $V \setminus \{v\}, \{v\}$ in G' form two cycle free components. Adding an edge between them does not create a cycle. Thus $(v, N[v])$ can be added to Q without creating a cycle. This contradicts our assumption that a maximal subset of edges Q, can have $|Q| < \frac{|V|}{2}$. Thus, $|Q| \geq \frac{|V|}{2}$.

2.b

Answer:

2.c

PAR-RANDOMIZED-CC can hook vertices to a root vertex such that it forms a star of depth at most 1. Also the expected number of edges thus hooked in a round is exactly $\frac{|E|}{2}$.
In RANDOM-HOOK during the second hooking phase, a vertex u hooked to root v can in turn add a new vertex to the star. Thus the depth of a start can be at most 2. Also the number of vertices hooked in a round is at least $\frac{|V|}{4}$.

2.d

Let X be the indicator random variable that represents whether L value of a vertex changes. For L value of a vertex to change, during the first phase of hooking, $C[v] = \text{HEAD}$ and $C[N[v]] = \text{HEAD}$.
The L value can change also during the second phase of hooking after flipping vertices that are not hooked. So the probability that L value of a vertex changes is at least equal to the probaility that L value changes during the first phase of hooking which is $\frac{1}{4}$.
Thus expected number vertices for which L value changes, $\mu = E[X] = |V| \times \frac{1}{4}$. $\delta = \frac{3}{4}$.
Applying, Chernoff Bound's lower tail:

$$P\left[X < \frac{|V|}{16}\right] < e^{\frac{\frac{|V|}{4} \times \frac{3}{4}^2}{2}} = e^{|V|/32}$$

Thus, $P\left[X \geq \frac{|V|}{16}\right] \geq 1 - \frac{1}{e^{|V|/32}}$.

2.e

Answer:

2.f

Some vertex will retain their PHD status, if they did not changed their $L[.]$ value in the RANDOM-HOOK. So, using **task 2.d**, probability that the vertices failed to change their $L[.]$ values,
$p_{fail_d} \leq \sum_{i=0}^{d} \frac{1}{e^{n_i/32}} \leq d \cdot \frac{1}{e^{n_d/32}} \leq \frac{d_{max}}{e^{n_{d_{max}}/32}} = \frac{\frac{1}{4}\log_{\frac{1}{\alpha}} n}{e^{\alpha^{2d_{max}} n/32}} = \frac{\frac{1}{4}\log_{\frac{1}{\alpha}} n}{e^{\alpha^{\frac{1}{4}\log_{\frac{1}{\alpha}} n}} n/32} = \frac{\frac{1}{4}\log_{\frac{1}{\alpha}} n}{e^{n^{-\frac{1}{2}} n/32}} = \frac{\log_{\frac{1}{\alpha}} n}{4e^{n^{\frac{1}{2}}/32}} =$
$o(n^{\frac{1}{c}})$ for some constant c

So, for each $d \in [0, d_{max}]$, $n_d \leq \alpha^{2d} \cdot n$ w.h.p. in n.

2.g

The question is wrong. Consider the complete graph as an input. At the beginning all edges and vertices are heavy. The first time you run random hook, the expected number of vertices that lose PhD status is at least $\frac{1}{4}n$. And so in the next level at least $\frac{1}{4}n(n-1) \nleq \alpha^d n$ heavy edges become light.

2.h

Because of RANDOM-HOOK,
Expected $n_d \leq (\frac{3}{4})^d n \leq \alpha^{2d} n$
Expected $n_{d_{max}} = O(n^{0.5})$
Therefore, expected number of edges $= O(n)$.

2.i

Let $n = |V|$ and $m = |E|$. PAR-RANDOMIZED-CC-3 is called $O(\log n)$ time. n_d and m_d are geometrically decreasing.
Expected $T_1 = O(n + m)$,
$T_p = n/p + m/p + \log p \log n)$. At the end, PAR-RANDOMIZED-CC-2 does not affect the asymptotic running time.

2.j

The timing info is provided in the \output\timing_info.txt

Input File	PR-CC-1	PR-CC-2	PR-CC-3
as-skitter-2j-CC-1-out.txt	28.359662	17.177	41.24
com-amazon-2j-CC-1-out.txt	2.891122	1.591265	3.17
com-dblp-2j-CC-1-out.txt	3.281750	1.7401	3.64
com-lj-2j-CC-1-out.txt	93.301557	68.3	136.84
com-orkut-pcc2.txt	254.32	231.422784	451.85
parallel_com-ca-2j-CC-1-out.txt	0.646374	0.655	1.247
roadNet-CA-2j-CC-1.txt	13.854537	5.323	15.99
roadNet-PA-2j-CC-1.txt	5.665667	2.84	8.46
roadNet-TX-2j-CC-1.txt	8.973965	3.57	10.525

Table 2: PR = Par-Randomized. Times are in seconds.

2.k

Answer: