ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

FS 2010 Prof. R. Wattenhofer / C. Lenzen

Principles of Distributed Computing
Exercise 10: Sample Solution

1 Diameter of the Augmented Grid

a) Each link connects to the target set with probability $p \in \Omega((\log^2 n)/n)$. Thus, for sufficiently large[1] n, the probability that $cn/\log n$ many links miss the set is bounded by

$$(1-p)^{cn/\log n} \leq e^{-pcn/\log n} \in e^{-\Omega(c \log n)} = \frac{1}{n^{\Omega(c)}}.$$

Now we exploit the power of the Big-O notation. Choosing a sufficiently large multiplicative constant in front of the $(cn/\log n)$-term, this becomes a bound of $1/n^c$, and choosing a large additive constant, we make sure that the bound holds also for the values of n that are not "sufficiently large". Thus, the probability that at least one link enters the set of $\Omega(\log^2 n)$ nodes is at least $1 - 1/n^c$, i.e., this event occurs w.h.p.

In order to obtain the same result using Chernoff's bound, let X_i, $i \in \{1, \dots, l\}$, where $l \in O(cn/\log n)$ is the number of considered links, be random variables that are 1 if the i^{th} link ends in the set (i.e., with the probability p from above) and 0 otherwise. Defining $X := \sum_{i=1}^{l} X_i$, we get that $\mathbb{E}[X] = pl$. Picking a constant $C > 0$ and properly adapting the constants in the $O(cn/\log n)$-term, we get that $\mathbb{E}[X] \geq C \log n$. Thus, Chernoff's bound states that for sufficiently large values of C and n (we need to cope with the fact that we do not know the constants in the O-term in Chernoff's bound), we have that $|X - \mathbb{E}[X]| < C \log n$ w.h.p. Because $P[X > 0] \geq P[|X - \mathbb{E}[X]| < \mathbb{E}[X]]$, this is what we are looking for.[2]

b) Because $|S| \in o(n)$, also $O(|S|) \subset o(n)$, i.e., the union of the set S ($|S|$ nodes) with the destinations of the $|S|$ random links and all grid neighbors of such nodes (at most $4|S|$ many nodes) has $o(n)$ nodes. Thus, always $(1-o(1))n$ nodes can be found which neither have been visited themselves nor have any neighbors that have been visited so far. Hence, regardless of the choice of the set S and any random links leaving S we have (sequentially) examined up to now, any uniformly independent random choice will contribute 5 new nodes with some probability $p \in 1-o(1)$. Now we use Chernoff's bound on the number of such "good" choices, yielding that it will be in $(1-o(1))|S|$ w.h.p. (instead of just in expectation).[3] Thus, in total we reach $(5 - o(1))|S|$ many nodes.

c) Recall that we may choose the constant c in "w.h.p." by ourselves. Thus, we may rule that in Chernoff's bound, it is $c' := c + 1$. Hence, the probability that in a given step our set grows by a factor $(5 - o(1))$ (provided that $|S| \in o(n)$, as we use part **b)**) is always at least $1 - 1/n^{c'}$. This means in at most a fraction of $1/n^{c'}$ of the events, something goes wrong in a single step. We need less than $\log n$ steps to get to $O(n/\log n)$ nodes, as the number of nodes more than quadruples in each step. In total, in a fraction of less than $\log n/n^{c'} = \log n/n \cdot 1/n^c < 1/n^c$ of all cases something goes wrong.

[1] This phrase means for some constant n_0, the statement will hold for all $n \geq n_0$.

[2] Small values of n are again dealt with by the additive constant in the O-notation. In general, it is always feasible to assume that n is "sufficiently large" when proving asymptotic statements.

[3] Since the expected value $\mathbb{E}[X]$ of the respective random variable X is large compared to $\log n$ (here we use $|S| \in \Omega((\log n)^2)$), the deviation from the expected value is with high probability in $o(\mathbb{E}[X])$.

d) Using the union bound again, we plug together the facts that (*i*) each node can reach $\Omega(\log^2 n)$ nodes following grid links only within $\log n$ steps, (*ii*) starting from these nodes, with high probability $O(n/\log n) \subset o(n)$ nodes can be reached within $O(\log n)$ more hops (part **b)**), (*iii*) from these nodes we reach with high probability the $(\log n)$-neighborhood (with respect to the grid) of any node (part **a**), and (*iv*) from there on we can reach the respective node with $\log n$ hops on the grid. Alltogether, with high probability in total $O(\log n)$ hops are necessary to reach some node v starting at some other node u. Finally, observe that we have $n(n-1) < n^2$ possible (ordered) combinations of nodes; choosing $c' := c + 2$ and applying the union bound once more, we infer that we have with high probability a path of length $O(\log n)$ between any pair of nodes, i.e., the diameter of the graph is in $O(\log n)$ w.h.p.

2 Greedy Routing in the Augmented Grid

The key observation is that the probability bound on advancing to the next phase is independent of the current node, as we always inspect fresh random links (recall that we get closer to the destination with each step). Assume for contradiction that it would take, say, $(C \log n)^2$ steps to reach the destination, where C and n are sufficiently large, with probability more than $1/n^c$. In each step, we proceed with probability $p \in \Omega(1/\log n)$ to the next phase. Let $X := \sum_{i=1}^{(C \log n)^2} X_i$, where the X_i are independent $0-1$ variables taking the value 1 with probability p. Thus, the probability to reach the destination is lower bounded by the probability that $X < \log n$.

The expected value $E[X]$ of X is at least $C \log n$ (here we use that C is chosen large enough to compensate for the multiplicative constant in the $\Omega((\log n)^2)$ term and n is sufficiently large that we can forget about additive constants). However, Chernoff's bound states that w.h.p., $|X - E[X]| \leq (C-1)\log n$ if C and n are also large enough to cope with the constants in the Big-O term in Chernoff's bound. Thus, our assumption that the probability to reach the destination within $(C \log n)^2 \in O(\log^2 n)$ many hops is smaller than $1 - 1/n^c$ must be wrong, implying that the desired statement holds indeed w.h.p.