BLACKBOARD PROOFS

CSE202 – WEEK 8

1. Hitting times

Here are the detailed proofs for slides 5 and 6. At step k, the random walk is at vertex X_k and moves to one of its neighbors with uniform probability. We write $u \in G$ to denote a vertex u in the graph G, $(u,v) \in G$ for an edge, $d(u)$ for the degree of the vertex u, n the number of vertices of G, m its number of edges and

$$T(u,v) := \mathbb{E}(\inf\{k \geq 1 \mid X_k = v,\ \text{with } X_0 = u\}), \qquad p_{uv} = \begin{cases} \frac{1}{d(u)}, & \text{if } (u,v) \in G \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 1.

$$\sum_{v \mid (u,v) \in G} T(v,u) = 2m - d(u).$$

Proof. (This is a step-by-step version of slide 6). Decomposing the expectation by the first step of the walk starting at a vertex w gives

$$T(w,u) = p_{wu} + \sum_{\substack{v \mid (w,v) \in G \\ v \neq u}} \frac{1}{d(w)}(1 + T(v,u))$$

since the first step can lead directly to u with probability p_{wu}, or to one of the other vertices v of w, from where the expected time to reach u is $T(v,u)$. The expression on the right-hand side simplifies in two ways: the sum over all $d(w)$ neighbors of w (including possibly u) of $1/d(w)$ is 1; the sum over all $v \neq u$ of the remaining part can be written as the sum over all v minus the term corresponding to $v = u$, which gives

$$T(w,u) = 1 + \frac{1}{d(w)} \sum_{v \mid (w,v) \in G} T(v,u) - p_{wu} T(u,u). \tag{1}$$

The next step is to multiply this identity by $d(w)$ and sum over all $w \in G$:

$$\sum_{w \in G} d(w) T(w,u) = \underbrace{\sum_{w \in G} d(w)}_{2m} + \underbrace{\sum_{w \in G} \sum_{v \mid (w,v) \in G} T(v,u)}_{\sum_{v \in G} d(v) T(v,u)} - \underbrace{\left(\sum_{w \in G} d(w) p_{wu} \right)}_{d(u)} T(u,u).$$

Three simplifications occur in the right-hand side: in the sum of $d(w)$ over all vertices $w \in G$, each edge of the graph is counted exactly twice, so that the first sum is $2m$; exchanging the order of summation in the second term shows that $T(v,u)$ occurs exactly $d(v)$ times in the sum; in the final sum, $d(w)p_{wu}$ is 1 when $(w,u) \in G$ and 0 otherwise, so that this sum is exactly $d(u)$.

The left-hand side is exactly equal to the middle term of the right-hand side (up to relabeling the index of the sum), so that in the end, we obtain

$$T(u,u) = \frac{2m}{d(u)}.$$

Now, using eq. (1) with $w = u$ and this value of $T(u,u)$ gives

$$\frac{2m}{d(u)} = 1 + \frac{1}{d(u)} \sum_{v|(u,v)\in G} T(v,u).$$

Multiplying by $d(u)$ concludes the proof. □

Each edge $(u,v) \in G$ occurs in a sum such as given by the lemma. As a consequence, $T(u,v) \le 2m-1$ (since $d(v) \ge 1$). Next, if $u \to u_1 \to \cdots \to v$ is a path of minimal length $\Delta(u,v)$ from u to v, summing these expectations over edges in the path gives the following.

Proposition 1. *For arbitrary vertices u and v in G, $T(u,v) \le (2m-1)\Delta(u,v)$.*

The final result of Slide 5 is the following.

Proposition 2. *The expected time $T(u,\cdot)$ for the walk to visit all nodes starting from u satisfies $T(u,\cdot) \le 2m(n-1)$.*

Proof. Consider a spanning tree $\mathcal{T}$ of G rooted at u. A depth-first traversal of $\mathcal{T}$ covers all vertices of G and thus gives an upper bound

$$T(u,\cdot) \le \sum_{(v,w)\in\mathcal{T}} (T(v,w) + T(w,v)),$$

since each edge of the tree is crossed once in each direction. This sum is bounded by

$$\sum_{(v,w)\in G} T(v,w) = \sum_{w\in G} \sum_{v|(v,w)\in G} T(v,w) = \sum_{w\in G} (2m - d(w)) = 2mn - 2m,$$

where the second equality is given by the lemma. □

2. Probability of ruin

The probability of reaching 0 starting from d when at each step the value is decreased by 1 with probability $1/k$ and increased by 1 otherwise satisfies

$$p(d) = \frac{1}{k} p(d-1) + \frac{k-1}{k} p(d+1),$$

with initial condition $p(0) = 1$ and the property that the limit for $d \to \infty$ is 0.

This linear recurrence of order 2 has two linearly independent solutions: the constant 1 and the sequence $(k-1)^{-d}$. Both can be checked by injecting into the recurrence. By linearity, all solutions to the recurrence are linear combinations of those two. The limit being 0 implies that the solution is a multiple of the second one only. Finally, the value at 0 shows that

$$p(d) = (k-1)^{-d}.$$

3. Lower bound on $\binom{3d}{d}$

The proof of
$$\binom{3d}{d} \geq \left(\frac{27}{4}\right)^d \frac{1}{3d+1}$$
can be done by induction on d. For $d = 0$, the inequality becomes an equality. Assuming the inequality to hold for d, it is sufficient to consider the ratio
$$\begin{aligned}\frac{\binom{3(d+1)}{d+1}}{\binom{3d}{d}} &= \frac{(3d+3)!}{(3d)!}\frac{(2d)!}{(2d+2)!}\frac{d!}{(d+1)!} = \frac{(3d+3)(3d+2)(3d+1)}{(2d+2)(2d+1)(d+1)} \\ &= \frac{27}{4}\frac{(d+2/3)(d+1/3)}{(d+1)(d+1/2)}.\end{aligned}$$
From there, it follows that
$$\begin{aligned}\binom{3(d+1)}{d+1} &\geq \left(\frac{27}{4}\right)^{d+1} \frac{(d+2/3)(d+1/3)}{(d+1)(d+1/2)}\frac{1}{3(d+1/3)} \\ &= \left(\frac{27}{4}\right)^{d+1} \frac{1}{3(d+4/3)}\underbrace{\frac{(d+4/3)(d+2/3)}{(d+1)(d+1/2)}}_{>1} \geq \left(\frac{27}{4}\right)^{d+1} \frac{1}{3d+4},\end{aligned}$$
concluding the induction.

4. Probability of success in $3n$ steps of Walkstat starting at distance d from a solution of 3-SAT

This is the detailed proof of Slide 17.

First the probability of success in at most $3n$ steps is at least the probability of success in at most $3d$ steps, which is at least the probability of success in exactly $3d$ steps. This happens for walks that perform $2d$ steps away from the solution and d steps towards it. Each of these $\binom{3d}{d}$ walks occurs with probability at least $(2/3)^d(1/3)^{2d}$ by the worst-case analysis of the previous slide. This gives a lower bound
$$\mathbb{P}(\text{success in } 3n \text{ steps starting from distance } d) \geq \binom{3d}{d}\left(\frac{2}{3}\right)^d\left(\frac{1}{3}\right)^{2d}.$$
Using the lower bound of the previous section and the inequality $d \leq n$ shows that this last quantity is lower bounded by
$$\frac{2^{-d}}{3n+1}.$$
Finally, among the 2^n possible assignments, $\binom{n}{d}$ are at distance d from the given solution A, so that the overall probability of success is obtained by summing these over d, leading to
$$\mathbb{P}(\text{success}) \geq 2^{-n}\sum_{d=0}^{n}\binom{n}{d}\frac{2^{-d}}{3n+1} = \frac{(3/4)^n}{3n+1},$$
where the last equality is simply the binomial theorem.