CSE 525 Randomized Algorithms & Probabilistic Analysis **Winter 2008**

Lecture 10: February 13

Lecturer: James R. Lee *Scribe: Noam Honda*

10.1 The power of two choices

As we discussed previously, if one throws n balls uniformly at random into n bins, then the maximum load is $\Theta(\frac{\log n}{\log \log n})$ with high probability. Now suppose that to throw a ball, we choose two bins uniformly at random and place the ball in the least loaded of the two choices. We will show that with high probability, the maximum load is only $\frac{\ln \ln n}{\ln 2} + O(1)$.

Let $\alpha_6 = \frac{n}{2e}$ and define $\alpha_{i+1} = \frac{e\alpha_i^2}{n}$ for $i \geq 6$. Let B_i be the number of bins with load $\geq i$ at the end of the process, and let $\mathcal{E}_i$ be the event that $\{B_i \leq \alpha_i\}$. Note that $\Pr[\mathcal{E}_6] = 1$ because $\frac{n}{2e} > \frac{n}{6}$.

Let $X(\beta)$ be a sum of n i.i.d. indicator random variables each with probability β of being 1 (i.e. X is distributed as $\text{Bin}(n, \beta)$, the binomial distribution with n independent trials and success probability β).

Claim 10.1 *Majorization:*

$$\Pr[B_{i+1} > \alpha_{i+1} \,|\, \mathcal{E}_i] \leq \Pr\left[X\left((\alpha_i/n)^2\right) > \alpha_{i+1}\right].$$

Proof: Define the *height of a ball* b to be the value h such that b is the hth ball thrown into its bin. We know that $B_{i+1} \geq \#\{\text{balls whose height is} \geq i+1\}$. Now for ball j, let Y_j be the indicator variable which is 1 precisely when balls j's height is $\geq i+1$. Let X_j be an indicator variable for choosing two bins of load $\geq i$ when the fraction of such bins is precisely $\frac{\alpha_i}{n}$. Observe that $Y_j \leq X_j$, so $\sum_{j=1}^n Y_j \leq \sum_{j=1}^n X_j$, where the X_j's are independent. Finally, $\sum_{j=1}^n X_j$ is distributed as $X((\alpha_i/n)^2)$. ■

Also, since $X(\beta)$ is a sum of i.i.d. $\{0,1\}$ random variables, we can apply a Chernoff bound:

Lemma 10.2 *For $\lambda \geq 1$, $\Pr[X(\beta) \geq \lambda e \cdot \beta n] \leq e^{-\lambda\beta n}$. In particular, $\Pr[X\left((\alpha_i/n)^2\right) > \alpha_{i+1}] \leq e^{-\alpha_i^2/n}$.*

Now, we have

$$\begin{aligned}
\Pr[\neg\mathcal{E}_{i+1}] &\leq \Pr[\neg\mathcal{E}_{i+1} \,|\, \mathcal{E}_i] + \Pr[\neg\mathcal{E}_i] \\
&= \Pr[B_{i+1} > \alpha_{i+1} \,|\, \mathcal{E}_i] + \Pr[\neg\mathcal{E}_i] \\
&\leq e^{-\alpha_i^2/n} + \Pr[\neg\mathcal{E}_i],
\end{aligned}$$

where we have used Claim 10.1 and Lemma 10.2 in the final line. Thus for $\alpha_i^2 > 2n \ln n$, we have $\Pr[\neg\mathcal{E}_{i+1}] \leq \frac{1}{n^2} + \Pr[\neg\mathcal{E}_i]$. We conclude (by induction) that for $\alpha_i^2 > 2n \ln n$, we have $\Pr[\mathcal{E}_{i+1}] \leq \frac{i+1}{n^2} \leq \frac{1}{n}$.

Now let $i^* = \min\{i : \alpha_{i^*}^2 < 2n \ln n\}$. Observe that $i^* = \frac{\ln \ln n}{\ln 2} + O(1)$, and by the preceding argument $\Pr[\neg\mathcal{E}_{i^*}] \leq \frac{1}{n}$.

Now we apply the Chernoff bound Lemma 10.2 and the majorization for one more step:

$$\Pr[B_{i^*+1} > 6 \ln n] \quad \leq \quad \Pr[B_{i^*+1} > 6 \ln n \,|\, \mathcal{E}_{i^*}] + \Pr[\neg\mathcal{E}_{i^*}]$$

$$
\begin{aligned}
&\leq \Pr\left[X(2(\ln n)/n) \geq 6\ln n\right] + \frac{1}{n} \\
&\leq e^{-6\ln n} + \frac{1}{n} \leq \frac{2}{n}.
\end{aligned}
$$

Finally, we use a simple union bound for the last step:

$$
\begin{aligned}
\Pr[B_{i^*+2} > 1] &\leq \Pr[B_{i^*+2} > 1 \mid B_{i^*+1} \leq 6\ln n] + \Pr[B_{i^*+1} > 6\ln n] \\
&\leq \Pr\left[X\left(\left\{\frac{6\ln n}{n}\right\}^2\right) > 1\right] + \frac{2}{n} \\
&\leq \left(\frac{6\ln n}{n}\right)^2 \cdot n + \frac{2}{n} = O\left(\frac{(\ln n)^2}{n}\right).
\end{aligned}
$$