HW #4 - Problem 3

Ryan St.Pierre (ras70)

October 4, 2017

Problem 3A

Every time the *insert* function iterates in attempts to find an empty space. There are two possible outcomes - either the space is free or it is empty.

$$\text{P(A[x] free)} = \frac{\text{free spaces}}{\text{total spaces}}\text{ , P(A[x] taken)} = \frac{\text{taken spaces}}{\text{total spaces}}$$

Let X be the number of times the *insert* function searches for a free block, that is the number of times $x = f(key)$ or $x = g(x)$ is run. In the case where X is geometrically distributed these probabilities of $A[x]$ being free and taken do not change every search iteration. However, in this case X is not exactly geometrically distributed. Every time the *insert* method fails the likelihood of a finding a free space increases and the probability of finding a taken space decreases. This is due to the fact that the g function returns a unique number every time it is called (given the number of calls is less than n). In other words, when a taken space is *checked* it is never checked again, and thus it is as though it does not exist.

Let i be the number of times the *insert* hash is called. The probability calculation for several values of i is given below:

i=1

This corresponds to the hash being successful on the first attempt.

$$P(X = 1) = 1 - \alpha$$

i=2

This corresponds to the hash being successful on the second attempt. For this to occur the hash has to fail the first time. The probability of failure on the first attempt is α. Knowing this location is a failure the probability of success on the next iteration is the probability of finding a free block where there are $n - 1$ total blocks. The number of free blocks has not changed and is still $n - \alpha n$. Therefore, the probability of success on the second attempt is $\frac{n-\alpha n}{n-1}$ or $1 - \frac{\alpha n-1}{n-1}$.

$$P(X = 2) = \alpha(1 - \frac{\alpha n - 1}{n - 1})$$

i=3

This corresponds to the hash being successful on the third attempt. For this to occur the hash has to fail the first two times. The first failure again has probability α, the second $\frac{\alpha n-1}{n-1}$.

$$P(X=3)=\alpha\frac{\alpha n-1}{n-1}(1-\frac{\alpha n-2}{n-2})$$

i=k

This expression can be generalized for an i value of k.

$$P(X=k)=\alpha\frac{\alpha n-1}{n-1}...\frac{\alpha n-(k-2)}{n-(k-2)}(1-\frac{\alpha n-(k-1)}{n-(k-1)})$$

Above we have shown the probability X takes a given value k. However, to calculate the expected value of X the following formula will be used,

$$\mathbb{E}[X]=\sum_{i=1}^{\infty}P[X\geq i]$$

Note: The above summation ranges from 1 to ∞ even though the size of the array is finite. It may be helpful to realize $P[X>n]=0$.

The probability that X is greater that or equal to i is the probability that the first $i-1$ hash attempts are failures. The probability that the first $i-1$ attempts are failures is

$$\alpha\frac{\alpha n-1}{n-1}...\frac{\alpha n-(i-2)}{n-(i-2)}$$

Therefore,

$$\mathbb{E}[X]=\alpha+\alpha\frac{\alpha n-1}{n-1}+\alpha\frac{\alpha n-1}{n-1}\frac{\alpha n-2}{n-2}+...$$

Noticing that each *individual* term in **each** product is less than or equal to α yields the following,

$$\begin{aligned}\mathbb{E}[X] &\leq \alpha+\alpha^2+\alpha^3+...\\ &\leq \textstyle\sum_{i=1}^{\infty}\alpha^{i-1}\end{aligned}$$

Let Y be a geometrically distributed random variable with probability $p = 1 - a$. Therefore,

$$\begin{aligned}
\mathbb{E}[Y] &= \textstyle\sum_{i=1}^{\infty} P[Y \geq i] \\
&= \textstyle\sum_{i=1}^{\infty} \alpha^{i-1} \\
&= \tfrac{1}{1-\alpha} && \text{from the homework hint}
\end{aligned}$$

Combining this known knowledge about $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ gives:

$$\begin{aligned}
\mathbb{E}[X] &\leq \textstyle\sum_{i=1}^{\infty} \alpha^{i-1} = \mathbb{E}[Y] = \tfrac{1}{1-\alpha} \\
\mathbb{E}[X] &\leq \tfrac{1}{1-\alpha}
\end{aligned}$$

It has been shown above that $\mathbb{E}[X]$ at most is $\frac{1}{1-\alpha}$.

Problem 3b

Again, let X be the number of times the *insert* function searches for a free block. Since X is a non-negative random variable Markov's inequality can be used. When $\lambda = 2$,

$$P[X \geq 2\,\mathbb{E}[X]] \leq \frac{1}{2}$$

Also, it is know $P[X \geq j]$ decreases as j increases. Since $\mathbb{E}[X] \leq \frac{1}{1-\alpha}$, $P[X \geq 2\,\mathbb{E}[X]] \geq P[X \geq 2\frac{1}{1-\alpha}]$. Therefore,

$$P[X \geq 2\frac{1}{1-\alpha}] \leq P[X \geq 2\,\mathbb{E}[X]] \leq \frac{1}{2}$$

This shows the following desired property:

$$P[X \geq 2\frac{1}{1-\alpha}] \ \leq \frac{1}{2}$$