Lecture 4, Thursday 12/4/01

Quicksort

- ```
  Quicksort(A,p,r)
    while p<r
      q=partition(A,p,r)
      quicksort(A,p,q-1)
      quicksort(A,q+1,r)
    end
  ```
- **To simplify, assume distinct elements.**
  - » **Lucky - always an even split:** $T(n) = 2T(n/2) + \Theta(n) \Rightarrow T(n) = \Theta(n \lg n)$
  - » **Unlucky:** $T(n) = T(0) + T(n-1) + \Theta(n) \Rightarrow T(n) = Q(n^2)$
- **How to avoid bad case ?**
  - » **Partitioning around middle element does not work!**
  - » **Idea: partition around a random element.**

42

Randomized Algorithms

- **Algorithm can “toss coins”.**
- **No specific input leads to worst-case behavior.**
- **Distinction between randomized algorithms and random data !**

43
  ```

Quick review of probability

- **Sample space S of "elementary events".**
  - » **Example: 36 ways of how 2 dice can fall.**
- **Event** $A \subseteq S$ **. Eg. "roll 3 with 2 dice".**
- **Probability distribution:** $P:\{A\} \to [0,1],\ 2^{|S|}\ values$
- **Properties:**

$$P(A) \geq 0,\ P(S) = 1$$
$$P(A \cup B) = P(A) + P(B)\ if\ \ A \cap B = \varnothing$$

44

Example

- **2 dice example:**

$$S = \{(1,1),(1,2),(2,1),\cdots,(6,6)\},\ |S| = 36$$
$$(5,6) \neq (6,5)!!$$
$$Event\ roll\ 4: \{(1,3),(2,2),(3,1)\}$$
$$\Pr[A] = \frac{|A|}{36} = \frac{3}{36}$$

- **Simple case of "inclusion/exclusion":**

$$\Pr[A \cup B] = P(A) + P(B) - P(A \cap B)$$
$$\leq P(A) + P(B)$$

45

Discrete Random Variable

- **Definition:**
  $X: S \to R$
  $event\ X=i \Leftrightarrow \{s \in S \mid X(s)=i\}$
  Ex: Uniform distr, 2 dice: $\Pr[X=5]=4/36$
- **Expected value:** $\sum_i i \Pr[X=i]$

SUM	Pr x 36	SUM x Pr x 36
1	0	0
2	1	2
3	2	6
4	3	12
………………………………………..		
12	1	12

		252

2 dice example

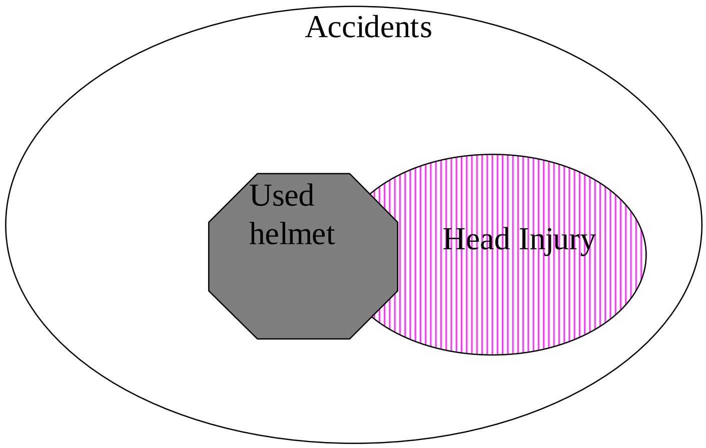
E[X]=252/36=7

Linearity of expectation

- **E[aX+bY]=aE[X]+bE[Y]**
- **Example: X - outcome of first, Y - outcome of second.**
  E[X]=E[Y]=[1+2+3+…+6]/6=3.5
  E[X+Y]=7, as before !
- **Independence:**

  $X \& Y\ independent\ iff\ \forall x,y: \Pr[X=i, Y=j]=\Pr[X=i]\Pr[Y=j]$
  $E[X \cdot Y]=E[X]E[Y]$

Conditional Probability



48

Conditional Probability

- **Definition:** $\Pr[X=i|Y=j]=\frac{\Pr[X=i,Y=j]}{\Pr[Y=j]}$

- **Conditional expectation:**

$$
\begin{aligned}
E_y\left[E_x[X|Y]\right] &= \sum_i \Pr[Y=i]E_x[X|Y=i] \\
&= \sum_i \Pr[Y=i]\sum_j j\Pr[X=j|Y=i] \\
&= \sum_i\sum_j j\Pr[X=j|Y=i]\Pr[Y=i] \\
&= \sum_j\sum_i j\Pr[X=j \cup Y=i]] \\
&= \sum_j j\Pr[\bigcup_i (X=j \cup Y=i)] \\
&= \sum_j j\Pr[X=j] \\
&= E[X]
\end{aligned}
$$

One of the most useful properties

49

Conditional expectation example

- **Consider 1-dice toss.**
- **Let X be result of the toss, and Y be the event that the result is above 2. (Y=1 if above 2, Y=0 otherwise.)**
- **Condition on Y.**
  Note that Pr[Y=0]=2/6, Pr[Y=1]=4/6.

$$E[X|Y=0]=\sum_i i\Pr[X=i|Y=0]=\frac{1}{6}\cdot\frac{1+2}{2/6}=\frac{3}{2}$$

$$E[X|Y=1]=\sum_i i\Pr[X=i|Y=1]=\frac{1}{6}\cdot\frac{3+4+5+6}{4/6}=\frac{9}{2}$$

$$E[X]=E[X|Y=0]\Pr[Y=0]+E[X|Y=1]\Pr[Y=1]$$

$$=\frac{3}{2}\cdot\frac{2}{6}+\frac{9}{2}\cdot\frac{4}{6}=3.5$$

50

Back to Quicksort

- **Partition around a randomly chosen element and let T(n) be the expected time to sort.**
- **Consider the case where the partition is (k, n-k-1).**
  In this case, the expected time to terminate is:

$$T(k)+T(n-1-k)+\Theta(n)$$

- **Condition on k being a specific value.**
  Note that any value of k, from 0 to n-1 is equally likely.

$$T(n)=\sum_k \Pr[(k,n-k-1)\text{ split}]\,T(n\,|\,(k,n-k-1)\text{ split})$$

$$=\frac{1}{n}\sum_k\left[T(k)+T(n-1-k)+\Theta(n)\right]$$

$$=\frac{2}{n}\sum_1^{n-1}\left[T(k)+\Theta(n)\right]$$

51

Solving the recurrence

We will try to prove that $T(n) \le an\lg n + b$

First, choose b large enough to satisfy: $T(1) \le a\lg 1 + b = b$

Inductive step:

$$T(n) = \frac{2}{n}\sum_{k=1}^{n-1} T(k) + \Theta(n) \le \frac{2}{n}\sum_{k=1}^{n-1}(ak\lg k + b) + \Theta(n)$$

$$= \frac{2}{n} a \underbrace{\sum_{k=1}^{n-1} k\lg k} + \frac{2}{n} nb + \Theta(n)$$

$$\le \frac{2}{n} a\left(\frac{1}{2}n^2\lg n - \frac{1}{8}n^2\right) + 2b + \Theta(n)$$

$$= an\lg n + b + \underbrace{(\Theta(n) + b - an/4)}_{\le 0 \text{ for large enough } a}$$

Need to prove that this is $\le \frac{1}{2}n^2\lg n - \frac{1}{8}n^2$

Note that using $\sum_{k=1}^{n-1} k\lg k \le n^2\lg n$ is not enough !!

Technical lemma

$n^2\lg n$ bound is trivial. Need a stronger bound

$$\sum_{k=1}^{n-1} k\lg k = \sum_{k=1}^{\lceil \frac{n}{2} - 1\rceil} k\lg k + \sum_{\lceil \frac{n}{2}\rceil}^{n-1} k\lg k$$

$\le \lg(n/2) = \lg n - 1$

$$\le \lg n \sum_{k=1}^{n-1} k - \sum_{k=1}^{\lceil \frac{n}{2} - 1\rceil} k$$

$$\le \lg n \frac{n(n-1)}{2} - \frac{(n/2-1)(n/2)}{2}$$

$$\le \frac{1}{2}n^2\lg n - \frac{n^2}{8}$$

HW: We proved O, now prove Ω.