COMPSCI 330 Lecture 16 Randomized Algorithms I

Wednesday, October 26, 2016 2:21 PM

- Recap: Probabilities, conditioning and expectations
- Randomized Quicksort
- Quick Selection

- idea: algorithm that tosses coins and make some decisions based on randomness.
- why? simpler algorithm, avoid bad cases

- random variable
  - variables that do not have fixed value
  - Example: $X = \begin{cases} 1 & \text{coin flip is head} \\ 0 & \text{coin tail} \end{cases}$ coin flip

    $Y = 1, 2, 3, \cdots, 6$ roll a die
  - $Pr[X=i]$: probability that $X=i$

    $Pr[X=0] = Pr[X=1] = \frac{1}{2}$

    $Pr[Y=1] = Pr[Y=2] = \cdots = Pr[Y=6] = \frac{1}{6}$
  - Example: Z: sum of points for 2 dice

    $Pr[Z=2] = \frac{1}{36}$, $Pr[Z=7] = \frac{1}{6}$
- Expectation

  similar to "average"

  $E[X] = \sum_i Pr[X=i] \cdot i$ $E[X] = \frac{1}{2}$ $E[Y] = 3.5$
  - linearity of expectation

    $E[X+Y] = E[X] + E[Y]$

    Y die Y' another die $Z = Y + Y'$

    $E[Z] = \sum_{i=2}^{12} Pr[Z=i] \cdot i$ (not very easy)

    $E[Z] = E[Y] + E[Y'] = 7$
- Conditioning

  $Pr[Z=i \mid Y=j]$: probability that $Z=i$, given we already know $Y=j$

  $Pr[Z=6 \mid Y=3]$: we know first die is 3, what is the probability

that sum of two dice is 6.

= same as probability that second die is 3

$$\Pr[Z=6 \mid Y=3] = \frac{1}{6} \qquad \Pr[Z=6] = \frac{5}{36}$$

$$\Pr[Z=2 \mid Y=3] = 0 \qquad \Pr[Z=2] = \frac{1}{36}$$

$$\Pr[Y'=3 \mid Y=4] = \Pr[Y'=3] = \frac{1}{6}$$

Y, Y' are independent.

- joint probability : two events happen simultaneously

$\Pr[Y=3, Y'=4]$: probability that first die is 3, second is 4

$$\Pr[Y=3, Y'=4] = \frac{1}{36} = \Pr[Y=3] \cdot \Pr[Y'=4 \mid Y=3]$$

$$\Pr[X=i, Y=j] = \Pr[X=i] \Pr[Y=j \mid X=i]$$

$$\Pr[Y=j \mid X=i] = \frac{\Pr[X=i, Y=j]}{\Pr[X=i]}$$

$$\Pr[Y=3, Z=4] = \frac{1}{36} = \Pr[Y=3] \cdot \Pr[Z=4 \mid Y=3]$$

$$\neq \Pr[Y=3] \cdot \Pr[Z=4]$$

$$\Pr[X_n = a_n, X_{n-1} = a_{n-1}, \dots, X_1 = a_1] = \Pr[X_1 = a_1] \cdot \Pr[X_2 = a_2 \mid X_1 = a_1]$$
$$\cdot \Pr[X_3 = a_3 \mid X_1 = a_1, X_2 = a_2]$$
$$\cdots$$
$$\cdot \Pr[X_n = a_n \mid X_1 = a_1, X_2 = a_2, \dots, X_{n-1} = a_{n-1}]$$

- $\mathbb{E}[Z \mid Y=i] = \sum_j \Pr[Z=j \mid Y=i] \cdot j$ (conditional expectation)

$$\mathbb{E}[Z \mid Y=2] = 5.5 = \mathbb{E}[Y \mid Y=2] + \mathbb{E}[Y' \mid Y=2]$$
$$= 2 + 3.5$$

- Law of total expectations

$$\mathbb{E}[Y] = \sum_i \Pr[X=i] \cdot \mathbb{E}[Y \mid X=i]$$

running time of a randomized algorithm

first random choice made by the algorithm

expected running time after we fix first random choice

- quicksort
  - pick a pivot
  - partition the array into smaller and larger parts

5 2 4 8 7 6 9 1

$\underline{5\ 2\ 4\ 1}$ 6 $\underline{8\ 7\ 9}$

- recurse on left and right parts.

- running time $O(n \log n)^*$ $*$: in worst case runnig time is $\Theta(n^2)$

- problem: if every time pivot is the smallest / largest, then quick sort takes $\Theta(n^2)$ time.

- idea: pick a random pivot number.

- how to analyze?

T_n : random variable: runnig time of randomized quick sort for array of size n.

want: $\mathbb{E}[T_n]$ is small.

$$T_n = \begin{cases} T_{n-1} + n & \text{if pivot is smallest} \quad Pr = \frac{1}{n} \\ T_1 + T_{n-2} + n & \text{pivot is second smallest} \quad Pr = \frac{1}{n} \\ \quad \cdots & \\ T_{i-1} + T_{n-i} + n & \text{pivot is } i\text{-th smallest} \quad Pr = \frac{1}{n} \end{cases}$$

$$\mathbb{E}[T_n] = \sum_{i=1}^{n} \underbrace{\frac{1}{n}}_{(Pr[\text{pivot} = i\text{th smallest}])} \cdot \underbrace{\mathbb{E}[T_{i-1} + T_{n-i} + n]}_{(\mathbb{E}[T_n \mid \text{pivot} = i\text{th smallest}])}$$

$$= n + \sum_{i=1}^{n} \left(\mathbb{E}[T_{i-1}] + \mathbb{E}[T_{n-i}] \right)$$

next time: will show $\mathbb{E}[T_n] \leq 4n \log_2 n$