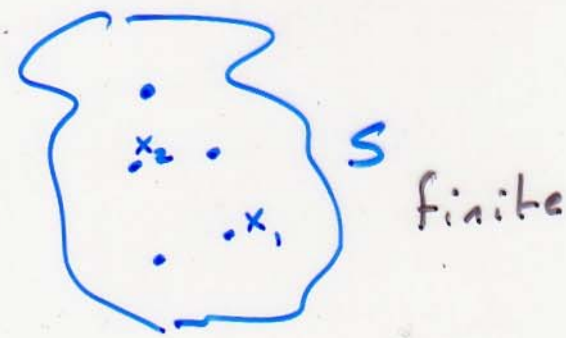
CS 223 Lec 24

Randomized Algorithms

- like a deterministic alg. except it can make "random" choices during execution.

Random Variables

X - discrete random variable
- drawn from a sample space S



- we also have a probability distribution

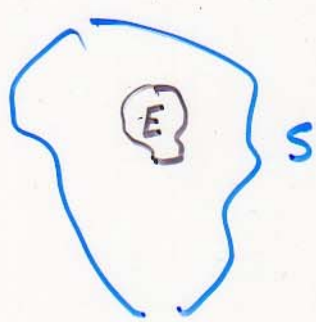
$$Pr(X = x) = \text{something}$$
$$(\text{for all } x \in S)$$

$$\sum_{x \in S} Pr(X = x) = 1$$

- An <u>event</u> is a subset of the sample space.

Let $E \subseteq S$

$$Pr(E) = \sum_{x \in E} Pr(X = x)$$



The Expectation of Random Variable

$$E(X) = \sum_{x \in S} x \cdot \Pr(X = x)$$

Example

$S = \{$ outcomes for flipping 10 fair coins $\}$

$X = \#$ of heads

$$E(X) = \sum_{x \in S} x \cdot \Pr(X = x)$$

$$= 0 \cdot \Pr(X=0) + 1 \cdot \Pr(X=1) + \cdots + 10 \cdot \Pr(X=10)$$

$$= 0 \cdot \left(\frac{1}{2}\right)^{10} + 1 \cdot \binom{10}{1} \left(\frac{1}{2}\right)^{10}$$

$$+ 2 \cdot \binom{10}{2} \left(\frac{1}{2}\right)^{10}$$

$$+ \cdots + \quad 10 \cdot \binom{10}{10} \left(\frac{1}{2}\right)^{10}$$

$$= \left(\frac{1}{2}\right)^{10} \Bigg(0 + 1 \cdot \binom{10}{1} + 2\binom{10}{2}$$

$$+ 3\binom{10}{3} + \cdots + 10\binom{10}{10} \Bigg)$$

$$= \cdots = 5$$

Linearity of Expectations

Let X, Y be random variables

Then $E(X + Y) = E(X) + E(Y)$

Let

$$X_i = \begin{cases} 1 & \text{if } i\text{th coin} = H \\ 0 & \text{otherwise} \end{cases}$$

<u>Claim</u>

$$X = X_1 + X_2 + \cdots + X_{10}$$

$$E(X) = E(X_1 + \cdots + X_{10})$$

$$= E(X_1) + E(X_2) + \cdots + E(X_{10})$$

$$E(X_i) = 0 \cdot Pr(X_i = 0) + 1 \cdot Pr(X_i = 1)$$

$$= Pr(X_i = 1) = 1/2$$

$$= \tfrac{1}{2} + \tfrac{1}{2} + \cdots \tfrac{1}{2} = 5.$$

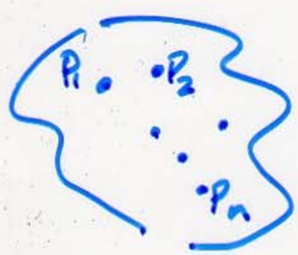
Indicator Variable.

Suppose E is an event

we define

$$I(E) = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0. & \text{otherwise.} \end{cases}$$

The Hiring Problem



candidates

Costs involved

C_i = interview cost

C_h = hiring cost

- keep interviewing candidates and hire a new candidate each time if he/she is better.

total cost =

$$n \cdot C_i + m \cdot C_h$$

m ↑ # of people hired.

<u>worst case</u>

$$\cancel{c_1 \; c_2}$$

$$P_1 < P_2 < P_3 < \ldots < P_n$$

hiring cost $= n \cdot C_h$

Idea to randomize the order of candidates

1st: find a random permutation of the numbers 1 to n.

$$\pi = \{51, 3, 6, 7, \ldots, 1, 5, 19, 30\}$$

2nd: interview the candidates in the order

$$P_{\pi(1)}, P_{\pi(2)}, P_{\pi(3)}, \ldots$$

Question: What is my expected hiring cost?

Let

X_i = indicate whether person $P_{\pi(i)}$ is hired

$$X = \sum X_i$$

What we want to find is $E(X)$

Need $E(X_i) = Pr(X_i = 1)$

$P_{\pi(1)}, P_{\pi(2)}, \ldots \boxed{P_{\pi(i)}}$

hire?

yes iff $P_{\pi(i)}$ is the best among the first i candidates

$$\Pr(X_i = 1) = \frac{1}{i}$$ ← because the best candidate is in the i^{th} positition with this probability.

So $E(X_i) = 1/i$

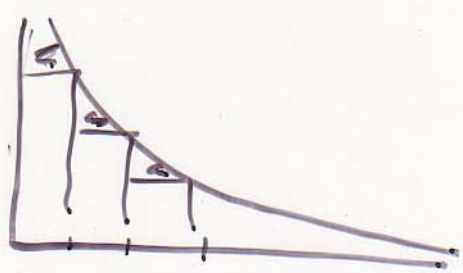
Let $X = X_1 + X_2 + \ldots + X_n$

$= $ total # of people hired

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

$$= 1/1 + 1/2 + 1/3 + \cdots + 1/n$$

$$= \sum_{i=1}^{n} 1/i = \ln n + O(1)$$



$$E(\text{hiring cost}) \le C_h \cdot (\ln n + O(1))$$

e.g. $n = 100$

$\ln_{100} = \underline{4}.xx.$

if we randomize first
only have to hire ~5
people.

Another example

<u>The</u> <u>Satisfiability</u> <u>Problem</u>

clause $(x_1 \underset{\text{OR}}{\vee} \neg x_5 \underset{\text{OR}}{\vee} x_{13})$

given a set of clauses

$\{C_1, C_2, \ldots, C_m\}$

problem is there a truth assignment to the Boolean variables, so that all clauses are satisfied (made true)

e.g. $C_1: x_1 \vee \neg x_2 \vee x_3$

$C_2: \neg x_1 \vee \neg x_2 \vee \neg x_3$

$C_3: \neg x_1 \vee x_2 \vee x_3$

set $\left.\begin{array}{l} x_1 = T \\ x_2 = T \\ x_3 = F \end{array}\right\} \checkmark$

Using Randomization for 2SAT.

this means every clause has exactly two literals

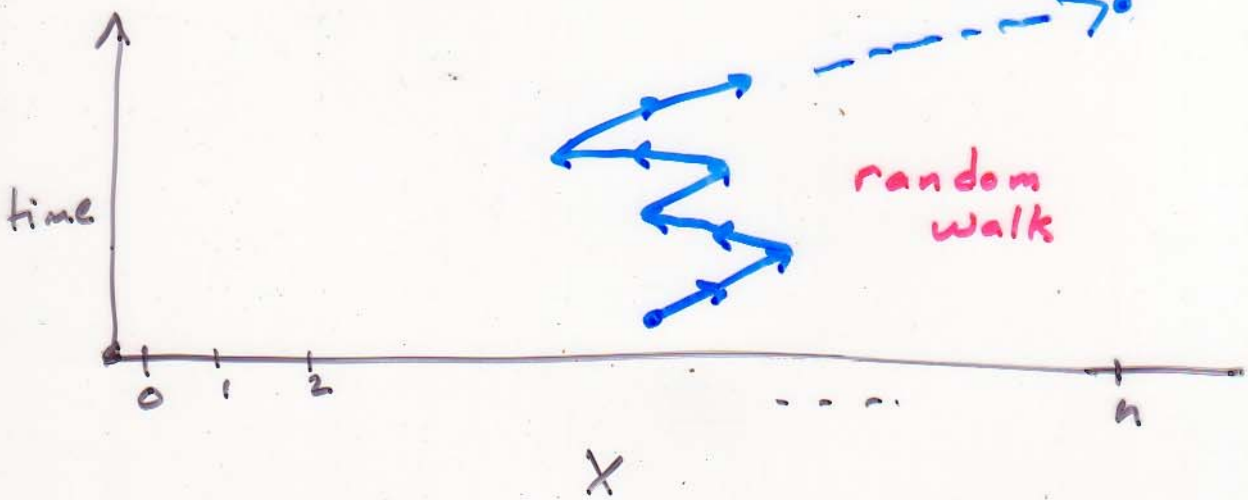
- start of with a random Boolean assignments
- while there is ≥ 1 unsatisfied clause & timer $< T$

{
 get an unsatisfied clause C_i
 randomly choose one of the variables in C_i and flip its truth assignment
}

- if all clauses satisfied, report "true"

- Let X = # of <u>correctly</u> assigned Boolean variables

$$X = \{0, 1, \ldots, n\}$$



at each flip,

either $X = X + 1$ Prob. $\geq 50\%$

or $X = X - 1$ Prob. $\leq 50\%$

<u>Fact</u> in $O(n^2)$ time, a random walk will reach n.