GIVEN: MC alg.

GOAL: Construct a LV alg. upon that

IDEA: repeat MC alg. until correct result is found

PROBLEM: need a (quick) check procedure

MC alg. has a runtime of $f(n)$

success probability of $p(n)$

checker $g(n)$

total expected runtime of LV algorithm:

$$\frac{f(n)+g(n)}{p(n)}$$

makes only sense if $g(n)$ does not take too much time

1.4. How to analyze rand. alg.?

Def.: (With High Probality) An event X occurs with high probability if with growing value of n the prob. of X converges to 1 or

$$P(X=x) \geq 1 - n^{-c}$$

for some constant c.

The event occurs with very high prob. if

$$P(X=x) \geq 1 - 2^{-cn}$$

$\frac{1}{2^a}$

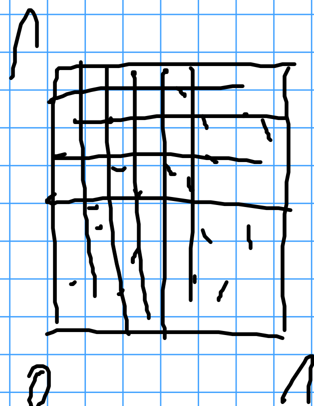
$a = c \log_2 n$

$\frac{1}{n^c} = n^{-c}$

$$P(X>t) \leq \frac{E(X)}{t}$$

$$P(X \leq t) = 1 - P(X>t) \geq \frac{E(X)}{t}$$

Example



unit square $[0,1]^2$
n points u.a.r.
boxes with side length $\log n / \sqrt{n}$

For $\varepsilon \in (0,1)$ a box is called ε-nice if it contains at least $(1-\varepsilon)\log^2 n$ points and at most $(1+\varepsilon)\log^2 n$ points

X - number of points in a fixed box

We would like to prove, that a fixed box being ε- nice happens w. high prob.

$$E(X) = \frac{n^2}{\frac{n^2}{\log^2 n}} = \log^2 n$$

in Markov Inequality

$$P(X > (1+\varepsilon) \cdot E(X)) \leq \frac{1}{1+\varepsilon}$$

Def.: (Chebyshev Inequality) Let X be a random variable, E(X) exp. value and σ the standard deviation. Then for all $t > 0$:

$$P(|X - E(X)| > t\sigma) \leq \frac{1}{t^2}$$

$$\left(\sigma = \sqrt{Var(X)} \qquad Var(X) = E(X^2) - E(X)^2 \right)$$

If $t = 2$, we can conclude that 75% the random variable falls within

$$[E(X) - 2\sigma, E(X) + 2\sigma]$$

Back to example:

$$P(|X - E(X)| > \underbrace{\varepsilon \cdot E(X)}_{= t \cdot \sigma}) \leq \frac{\sigma^2}{\varepsilon^2 \cdot E(X)^2}$$

$$t = \frac{\varepsilon \cdot E(X)}{\sigma}$$

$$\left(Var(X) = n \cdot p \cdot q = \cancel{n} \cdot \frac{\log^2 n}{\cancel{n}} \cdot \left(1 - \frac{\log^2 n}{n}\right) \right)$$

$$P(|X - E(X)| > \varepsilon \cdot E(X)) \leq \frac{\log^2 n \left(1 - \frac{\log^2 n}{n}\right)}{\varepsilon^2 \left(\log^2 n\right)^2}$$

$$= \frac{n - \log^2 n}{\varepsilon^2 n \log^2 n}$$

For $n \to \infty$ the term goes to zero, i.e. the box being ε-nice happens with high probability.

1.2. Analyzing LV

<u>Def.</u>: <u>Completeness</u>

If it is guaranteed to finish in t_{max} steps (instance-dependent) a LV alg. is called to be complete.
If the prop. to solve the problem converges to 1 with increasing runtime, the LV alg. is called app. complete.

Find Repeated Elem. example:

LV might run forever

$P(\text{success in } k \text{ rounds}) = 1 - \left(\frac{3}{4}\right)^k$

for $k \to \infty$ this term converges to 1

↳ app. complete

Detect Drug Example:

After at most $\frac{n}{2} + 1$ steps the LV alg. is done

↳ complete

2. Analyze MC

- one-sided and two-sided errors
- example for two-sided error:

10.000 light bulbs. Are there at least 80% working?

⇒ MC. alg.: test 10, if less than 8 work, return 'no', else return 'yes'

Prove: • $E\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k} E(X_i)$

• $P(X > t) \leq \frac{E(X)}{t}$

$$P(X>t) > \frac{E(X)}{t} \Rightarrow t\cdot P(X>t) > E(X)$$

$$E(X) = \sum_{i=1}^{k} P(X=x_i)\cdot x_i$$

$$= \sum_{i=1}^{l} P(X=x_i)\cdot x_i + \sum_{i=l+1}^{k} P(X=x_i)\cdot \underline{x_i}$$

$$\sum_{x_i \le t} \qquad\qquad \sum_{x_i > t}$$

$$\geq \sum_{x_i\le t} P(X=x_i)\cdot x_i + t\cdot \underbrace{\sum_{x_i>t} P(X=x_i)}_{= P(X>t)} \cancel{=P(X>t)}$$

$$> \sum_{x_i\le t} P(X=x_i)\cdot x + t\cdot \frac{E(X)}{t}$$

$$= \underbrace{\quad '' \quad}_{\geq 0 \text{ because } X\geq 0} + E(X) \quad ↯$$

$$E\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k} E(X_i)$$

Prove

$$E(X) = \sum_{i=1}^{n} x_i \cdot p_i \quad \text{— } P(X = x_i)$$

$$E\left(\sum_{i=1}^{k} X_a\right) = \sum_{i=1}^{n} (X_{1i} + \ldots + X_{ki}) \cdot P_i$$

$$= \sum_{j=1}^{k} \sum_{i=1}^{n} X_{ji} \cdot p_i \quad \text{— } E(X_j)$$

$$= \sum_{j=1}^{k} E(X_j) \qquad \square$$

---

2 Rand. Alg. & Data Structures

2.1. The Max 3-SAT Problem

Def.: Given a set of clauses $C_1, \ldots, C_k$ (each of length 3) over a set of boolean variables $X = \{x_1, \ldots, x_n\}$

(Example: $(x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee x_4 \vee \neg x_5)$)
$k = 2, n = 5$

Goal is to find an assignment for X which satisfies as many clauses as possible.

Simple rand. alg.: for each x_i assign true or false randomly, each with prob. $1/2$

Lemma: The expected number of satisfied clauses is $\geq \frac{7}{8} q$.

Proof: Let Y denote rand. var. which counts satisfied clauses. $E(Y) = E\left(\sum_{i=1}^{q} Y_i\right)$

Y_i being 1 if clause C_i is satisfied
0 otherwise

So $E(Y_i) = P(C_i \text{ is satisfied})$

C_i is not satisfied only if all three variables got the wrong assignment.

Prob. for that $\left(\frac{1}{2}\right)^3 = \frac{1}{8}$

Thus $E(Y_i) = \frac{7}{8}$.

$E(Y) = \sum_{i=1}^{q} E(Y_i) = q \cdot \frac{7}{8} = \frac{7q}{8}$ //

Lemma: For every instance of the Max 3-SAT problem there exists an assignment satisfying at least a $\frac{7}{8}$ fraction of all clauses.

$\Rightarrow$ non-constructive existence proof
$\Leftarrow$ prob. method

Exercise. Prove that for every instance of Max3Sat with $k \leq 7$ all clauses can be satisfied

LV: • run MC alg. with random assignment $(O(n))$
• check if at least $\frac{7}{8}k$ clauses are satisfied $(O(k))$

• if not enough clauses are fulfilled run MC again

What about exp. runtime?

$p_j = P(Y=j)$, $j = 0, \dots, k$ prob. to satisfy exactly j clauses

Success Probability

$$p = \sum_{j \geq \frac{7}{8}k} p_j$$

$$\frac{7k}{8} = E(X) = \sum_{j=0}^{k} j \cdot p_j$$

$$\frac{7}{8}k = \sum_{j < \frac{7}{8}k} j \cdot p_j + \sum_{j \geq \frac{7}{8}k} j\, p_j$$

We define k' to be the largest natural number smaller than $\frac{7}{8}k$ (if $\frac{7}{8}k \in \mathbb{N}$ $k' = \frac{7}{8}k - 1$, else $k' = \lfloor \frac{7}{8}k \rfloor$)

On that basis:

$$\frac{7}{8}k \leq \sum_{j < \frac{7}{8}k} k' p_j + \sum_{j > \frac{7}{8}k} k \cdot p_j = k' \sum_{j < \frac{7}{8}k} p_j + k \cdot \sum_{j \geq \frac{7}{8}k} p_j$$

We further observe that $\sum_{j < \frac{7}{8}k} p_j = 1 - p$ and so

$$\frac{7}{8}k \leq k'(1-p) + k \cdot p \leq k' + kp$$

in terms of p this equals:

$$p \geq \frac{\frac{7}{8}k - k'}{k}$$

for k': $\frac{7}{8}k - k' \geq \frac{1}{8}$

$$\Rightarrow \quad p \geq \frac{1}{8k}$$

Theorem: The LV alg. for Max 3-SAT satisfies at least $\frac{7}{8}$ of the clauses in expected $8k$ rounds, each costing us $O(n \cdot k)$.

In total we have a runtime of $O(nk + k^2)$ which is polynomial.