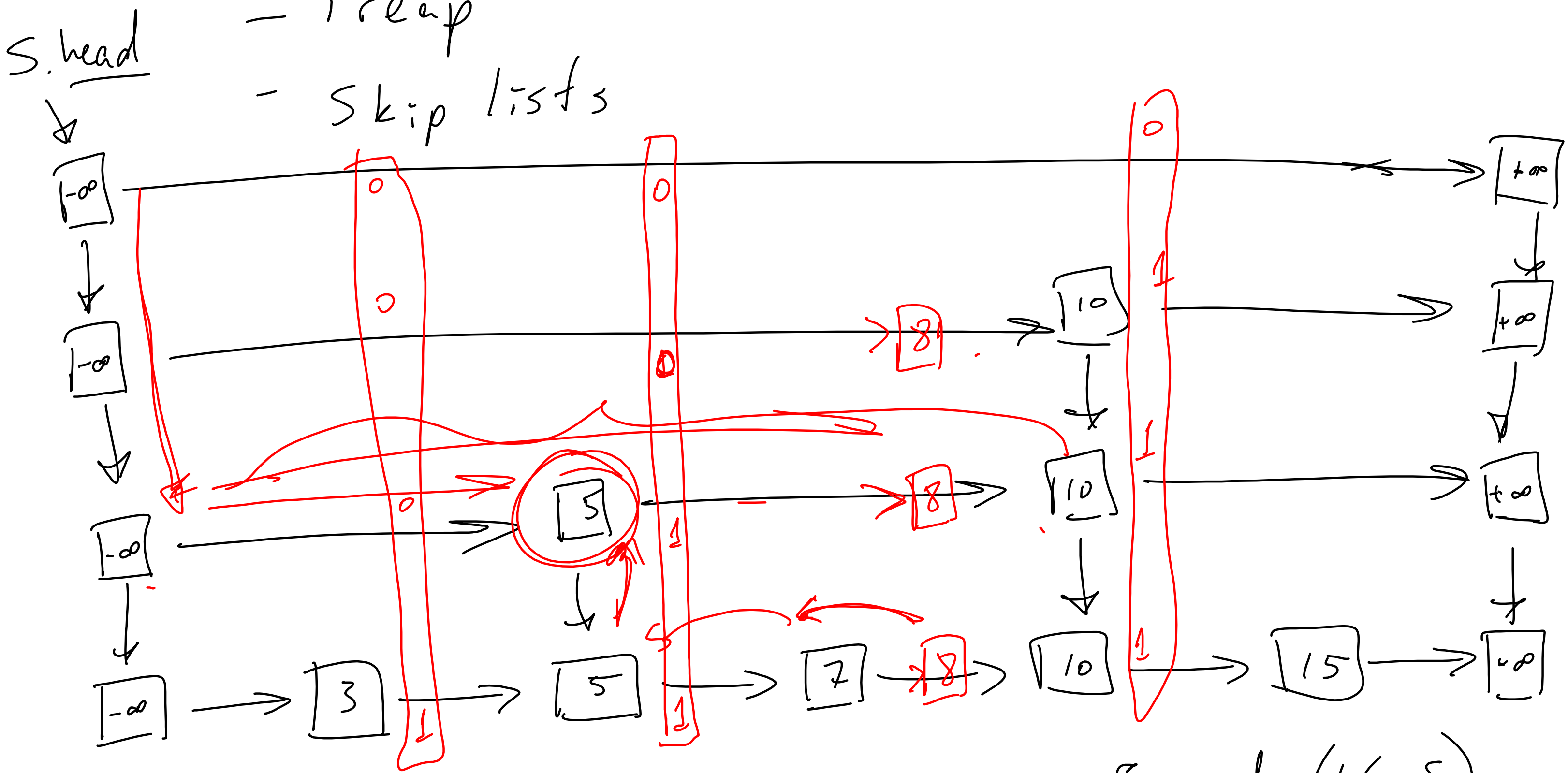
Last week:

- Treap
- Skip lists



Search(16, S)

Predecessor-Search(k, S)

```
v = S.head
while v.down ≠ nil
    while (v.next ≠ nil & v.next.key ≤ k)
        v = v.next
    v = v.down
while (v.next ≠ nil & v.next.key ≤ k)
    v = v.next
return v
```

Analysis of (Randomized) Skip List Height

Lemma: The height $h(T)$ of the skip list T on n keys is at most $(c+1)\log n$ w/ prob. $1 - \frac{1}{n^c}$

$$\Pr[h(T) < (c+1)\log n] \geq 1 - \frac{1}{n^c}$$

for any constant $c > 0$

high probability

Proof: A node v will at most at level k if k heads came up in a row.

$$\Pr[h(v) \geq k] = \Pr[\text{first } k \text{ flips are heads}] = \frac{1}{2^k}$$

$$\Pr[h(T) \geq (c+1)\log n] = \Pr\left[\max_{v \in T} h(v) \geq (c+1)\log n\right]$$

$$= \Pr[h(v_1) \geq (c+1)\log n \text{ or } h(v_2) \geq (c+1)\log n \text{ or } \dots h(v_n) \geq (c+1)\log n$$

$$\le \Pr[h(v_1) \ge (c+1)\log n] +$$

$$\Pr[h(v_2) \ge (c+1)\log n] +$$

$$\vdots$$

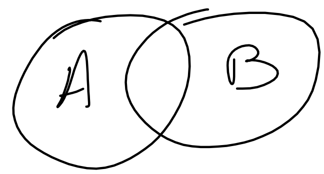
$$\Pr[h(v_n) \ge (c+1)\log n]$$

$$\le \sum_{i=1}^{n} \frac{1}{2^{(c+1)\log n}} = \frac{n}{\left(2^{\log n}\right)^{(c+1)}}$$

$$= \frac{n}{n^{(c+1)}} = \frac{1}{n^{(c+1)-1}} = \frac{1}{n^{c}}$$

Union Bound

$$\Pr[A \text{ or } B] \le \Pr[A] + \Pr[B]$$



$$|A \cup B| \le |A| + |B|$$

$$\Pr[h(T) < (c+1)\log n] \ge 1 - \Pr[h(T) \ge (c+1)\log n]$$

$h(T) \in O(\log n)$

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

with high probability

$$1 - \frac{1}{n}$$

$$\Pr[h(T) = O(\log n)] = 1 - \frac{1}{n^{O(1)}}$$

Insert (k, S)

$v \leftarrow$ search (k, S)

if $v.key \neq k$

$v' =$ new Node (k)

ListInsert (v, v')

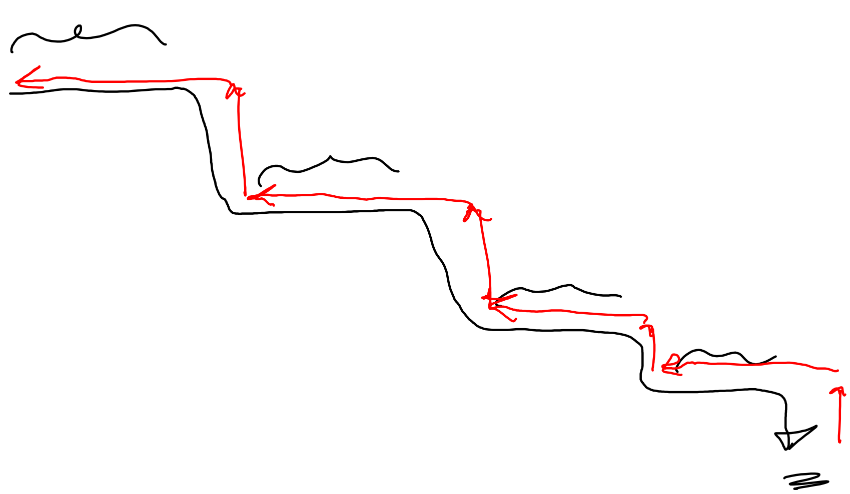
bit = coinFlip()

while (bit == 1)

promote a copy of v' to the list above

bit = coinFlip()

(Randomized) Skip List Search Runtime Analysis



Backward-Search (v, S)

while $(v \neq S.\text{head})$

if $v.\text{up} \neq \text{nil}$

$v = v.\text{up}$

else

$v = v.\text{left}$

if canFlip() = 1

$v = v.\text{up}$

else

$v = v.\text{left}$

Expected # of tails is equal to expected # of heads

$$E[\text{runtime(search)}] = 2 \cdot E[\#\text{of heads}]$$

$$= 2 \cdot E[\text{height of skip list}]$$

Analysis of Expected height of Skip List

For any positive integer random variable X

$$E[X] = \sum_{x} x \cdot Pr[X = x]$$

$$\boxed{E[X] = \sum_{x=1}^{\infty} Pr[X \geq x]}$$

Lemma:

$$\sum_{x=1}^{n} x \cdot f(x) = \sum_{x=1}^{n} \sum_{i=x}^{n} f(i)$$

$$\sum_{x=1}^{n} x \cdot f(x) = 1 \cdot f(1) + 2 \cdot f(2) + 3 \cdot f(3) + \dots + n \cdot f(n)$$

$$= (f(1) + f(2) + \dots f(n)) + (1 \cdot f(2) + 2 \cdot f(3) + \dots + (n-1) f(n))$$

$$= (f(1) + f(2) + \dots + f(n)) + (f(2) + f(3) + \dots + f(n)) + (f(3) + f(4) + \dots + f(n)) + \dots + (f(n))$$

$$\sum_{x=1}^{n} \boxed{\sum_{i=x}^{n} f(i)}$$

Lemma: For any non-negative integer random variable X

$$E[X] = \sum_{x=1}^{\infty} Pr[X \geq x]$$

$f(x) = Pr[X = x]$

Proof:

$$E[X] = \sum_{x=1}^{\infty} x \cdot \underbrace{Pr[X = x]}_{f(x)} = \sum_{x=1}^{\infty} \boxed{\sum_{i=x}^{\infty} \underbrace{Pr[X = x]}_{f(x)}}$$

$$= \sum_{x=1}^{\infty} Pr[X \geq x]$$

$$\sum_{i=x}^{\infty} Pr[X = x] = Pr[X \geq x]$$

$$\| $$

$$Pr[X = x] + Pr[X = x+1] + \dots + Pr[X = \infty]$$

Expected height of Skip List

$$E[height(S)] = E[\max_v h(v)]$$

Let $H = \max_v h(v)$

$$E[H] = \sum_{x=1}^{\infty} Pr[H \geq x] =$$

$$= \underbrace{\sum_{x=1}^{\log n} Pr[H \geq x]} + \sum_{x=(\log n)+1}^{\infty} Pr[H \geq x]$$

$$\leq \sum_{x=1}^{\log n} 1 + \sum_{x=\log n+1}^{\infty} Pr[H \geq x]$$

$$= \log n + \underline{\sum_{x=\log n+1}^{\infty} Pr[H \geq x]}$$

$$Pr[H \geq x] = Pr[\underline{h(v_1) \geq x} \text{ or } \underline{h(v_2) \geq x} \text{ or } \dots \underline{h(v_n) \geq x}]$$

by union bound

$$\leq \sum_{i=1}^{n} Pr[\underline{h(v_i) \geq x}]$$

$$\leq n \cdot \frac{1}{2^x}$$

$$\sum_{x=\log n+1}^{\infty} Pr[H \geq x] \leq \sum_{\underline{x=\log n+1}}^{\infty} n \cdot \frac{1}{2^x} = n \cdot \sum_{\underline{i=1}}^{\infty} \frac{1}{2^{\log n+i}}$$

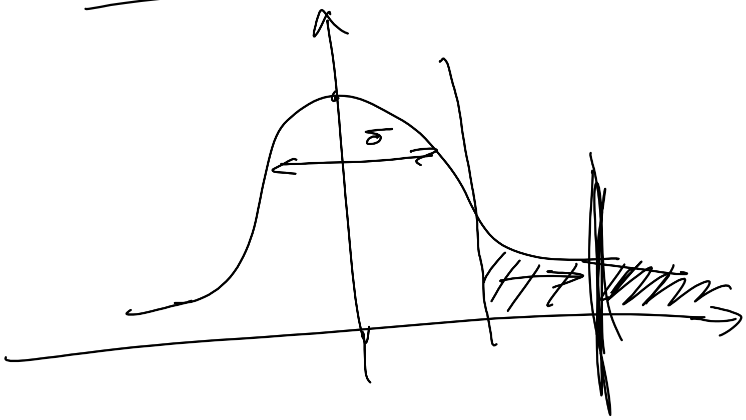
$$= \cancel{n} \cdot \sum_{i=1}^{\infty} \frac{1}{\cancel{n} \cdot 2^i} = \sum_{i=1}^{\infty} \frac{1}{2^i} = \underline{\underline{1}}$$

$$E[H] = \log n + \sum_{x=\log n+1}^{\infty} \Pr[H \geq x]$$

$$= \underline{\underline{\log n + \underline{1}}}$$

$$h \underline{\underline{<}} c \cdot \log n \quad \text{w prob} \leq \frac{1}{n^{c-1}}$$

Tail Bounds



Markov's Inequality

Lemma: Non-negative random variable X

$$\Pr[X \geq a] \leq \frac{E[X]}{a}$$

Proof:

$$E[X] = \sum_{x=0}^{a-1} x \cdot \Pr\{X = x\} + \sum_{x=a}^{\infty} x \cdot \Pr[X = x]$$

$$\geq \sum_{x=a}^{\infty} x \cdot \Pr[X = x]$$

$$\geq a \cdot \sum_{x=a}^{\infty} \Pr[X = x]$$

$$= a \cdot \Pr[X \geq a]$$

$$\Rightarrow \Pr[X \geq a] \leq \frac{E[X]}{a}$$

The Height of Treaps

For what values of h

$$\Pr[\text{Height}(\text{Treap}) \geq h] \leq \frac{1}{n}$$

$$\Pr[h(v_1) \geq h \text{ or } h(v_2) \geq h \text{ or } \dots \text{ or } h(v_n) \geq h] \leq$$

$$\leq \Pr[h(v_1) \geq h] + \Pr[h(v_2) \geq h] + \dots + \Pr[h(v_n) \geq h]$$

$$= n \cdot \Pr[h(v_k) \geq h]$$

Consider the k-th smallest key node v_k

$$n \cdot \Pr[h(v_k) \geq h] \leq \frac{1}{n}$$

$$\Rightarrow \Pr[h(v_k) \geq h] \leq \frac{1}{n^2}$$

By Markov's inequality:

$$\Pr[h(v_k) \geq h] \leq \frac{E[h(v_k)]}{h} = \frac{\log n}{h} \leq \frac{1}{n^2}$$

$$\Rightarrow h \geq n^2 \log n$$

$$\Pr[\text{height of treap} \geq \underline{n^2 \log n}] \leq \frac{1}{n}$$

Chebyshev's inequality

Def'n: Two random variables X & Y are independent iff

$$\Pr[X = x \;\&\; Y = y] = \Pr[X = x] \cdot \Pr[Y = y]$$

$$\Pr[X = x \mid Y = y] = \Pr[X = x]$$

Def'n: A set of random variables $X_1, X_2, \ldots, X_n$ are k-wise independent if every subset of k variables are independent

Example; X & Y are independent binary random variables

$$Z = (X + Y) \bmod 2$$

X, Y, Z are pair-wise independent but not 3-wise independent

Corollary: if $X_1, \ldots, X_n$ are k-wise independent, then they are also k'-wise for any $2 \le k' \le k$

Theorem (Chebyshev's inequality):

Let $X_1, X_2, \dots, X_n$ be pairwise independent non-negative indicator random variables, and

$$X = \sum_{i=1}^{n} X_i \quad \text{with} \quad E[X] = \mu$$

then

$$\Pr[(X-\mu)^2 \geq a] < \frac{\mu}{a}$$

for any $a > 0$.

$k^2 = a$

$\Pr[(X-\mu)^2 \geq k^2]$

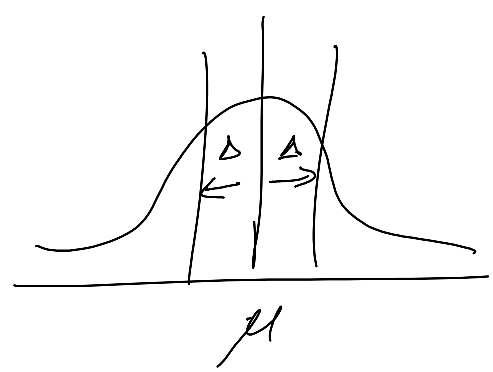
$= \Pr[|X-\mu| \geq k]$

Alternatively:

$$\Pr[|X-\mu| \geq k] < \frac{\mu}{k^2}$$

Let $\Delta = k$:

Additive bounds:



$$\Pr[X \geq \mu + \Delta] \leq \frac{\mu}{\Delta^2}$$

$$\Pr[X \leq \mu - \Delta] \leq \frac{\mu}{\Delta^2}$$

Let $\delta\mu = k$

multiplicative bounds:

$$\Pr[X \geq \underline{\underline{(1+\delta)}}\mu] \leq \frac{1}{\delta^2\mu}$$

$$\Pr[X \leq (1-\delta)\mu] \leq \frac{1}{\delta^2\mu}$$

Chernoff Bound

Theorem:

If indicator random variables $X_1 \ldots X_n$ are fully independent, and $X = \sum_{i=1}^{n} X_i$ then

$$\Pr[X \geq a] \leq e^{a-\mu}\left(\frac{\mu}{a}\right)^{a}$$

for all $a \geq \mu = E[X]$

Additive bounds

$$\Pr[X \geq \mu + \Delta] \leq e^{\Delta}\left(\frac{\mu}{\mu+\Delta}\right)^{\mu+\Delta}$$

$$\Pr[X \leq \mu - \Delta] \leq e^{-\Delta}\left(\frac{\mu}{\mu-\Delta}\right)^{\mu-\Delta}$$

Multiplicative Bounds

$$\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$$

$$\Pr[X \leq (1-\delta)\cdot\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$$

If $0 < \delta < 1$:

$$\Pr\left[X \geq (1+\delta)\mu\right] \leq e^{-\frac{\delta^2 \mu}{3}}$$

$$\Pr\left[X \leq (1-\delta)\mu\right] \leq e^{-\frac{\delta^2 \mu}{2}}$$

$\boxed{\text{Examples}}$

Flip a fair coin N times.

How many heads do I expect?

$$X_i = \begin{cases} 0 & \text{if } i\text{th flip is tails} \\ 1 & \text{if } i\text{th flip is heads} \end{cases}$$

$$\mu = E[X_i] = \Pr[X_i = 1] = \frac{1}{2}$$

X = total # of heads

$$X = \sum_{i=1}^{N} X_i \Rightarrow E[X] = E\left[\sum X_i\right] = \sum_{i=1}^{N} E[X_i] = N \cdot \frac{1}{2} = \frac{N}{2}$$

Q: What's the probability that # of heads is $\geq \frac{3}{4}N$?

$\geq \alpha N \quad \frac{1}{2} < \alpha < 1$

<u>Markov's inequality:</u>

$$\Pr[X \geq a] \leq \frac{\mu}{a}$$

$$\Pr[X \geq \alpha N] \leq \frac{N/2}{\alpha N} = \frac{1}{2\alpha}$$

E.g. $\alpha = \frac{3}{4}$: $\Pr[X \geq \frac{3}{4}N] \leq \frac{1}{2 \cdot \frac{3}{4}}$

$$\leq \frac{2}{3}$$

<u>Chebyshev's inequality</u>

<u>Additive:</u> $\Pr[X \geq \mu + \underline{\Delta}] \leq \frac{\mu}{\Delta^2}$

$$\Pr[X \geq \alpha N] = \Pr[X \geq \tfrac{1}{2}N + \underbrace{(\alpha - \tfrac{1}{2})N}_{\Delta}]$$

$$\leq \frac{N/2}{(\alpha - \frac{1}{2})^2 N^2} = \frac{1}{2(\alpha - \frac{1}{2})^2 N}$$

$\alpha = \frac{3}{4}$: $\Pr\left[X \geq \frac{3}{4}N\right] \leq \frac{1}{2\left(\frac{3}{4}-\frac{1}{2}\right)^2 N} = \frac{8}{N}$

Multiplicative: $\Pr\left[X \geq (1+\delta)\mu\right] \leq \frac{1}{\delta^2 \mu}$

$\Pr\left[X \geq \alpha N\right] \leq \frac{1}{\delta^2 \mu}$

$\frac{(1+\delta)}{2} \cdot \cancel{N} = \alpha \cancel{N}$

$\boxed{\delta = 2\alpha - 1}$

$= \frac{1}{(2\alpha-1)^2 \cdot \frac{N}{2}} = \frac{2}{(2\alpha-1)^2 N}$

$\frac{1}{2} < \alpha \leq 1$

$0 < \delta = 2\alpha - 1 \leq 1$

$\alpha = \frac{3}{4}$: $\Pr\left[X \geq \frac{3}{4}N\right] = \frac{2}{\left(2 \cdot \frac{3}{4} - 1\right)^2 N}$

$= \frac{2}{\left(\frac{1}{2}\right)^2 N} = \frac{8}{N}$

Chernoff Bound

$0 < \delta = 2\alpha - 1 < 1$

$\Pr\left[X \geq (1+\delta)\mu\right] \leq e^{-\frac{\delta^2 \mu}{3}}$

$\Pr\left[X \geq \alpha N\right] \leq e^{-\frac{(2\alpha-1)^2 \cdot \frac{N}{2}}{3}} = e^{-\frac{(2\alpha-1)^2 N}{6}}$

$\alpha = \frac{3}{4}$: $\Pr\left[X \geq \frac{3}{4}N\right] \leq e^{-\frac{\left(2 \cdot \frac{3}{4} - 1\right)^2 \cdot N}{6}} = e^{-\frac{1}{4} N / 6}$

$$= e^{-\frac{N}{24}} = \frac{1}{e^{N/24}}$$

Markov: $\Pr[X \geq \frac{3}{4}N] \leq \frac{2}{3}$ ←

Chebyshev: $\Pr[X \geq \frac{3}{4}N] \leq \frac{8}{N}$ ←

<u>Chernoff</u>: $\Pr[X \geq \frac{3}{4}N] \leq \frac{1}{e^{N/24}}$ ←

# of heads