ICS 621: Analysis of Algorithms	Fall 2021
Lecture 8	
Prof. Nodari Sitchinava	*Scribe: In Woo Park, Darlene Agbayani, Michael Rogers*

1 Overview

Previously, we introduced Treaps, defining their basic structure. As well as Treap Operations, including: Treap-Insert(T, v), Treap-Delete(v), Treap-Increase-Priority$(v, new_p riority)$, Treap-Split(T, key) and Treap-Merge$(T_<, T_>)$. After defining the Treap Operations, we analyzed their runtimes. We also introduced Skip Lists and define their structure. Then further discussed Skip Lists in context of Intuition and Insertion Operation.

In this lecture we will continue discussing the topic of randomized search trees, beginning with a quick recap of what we learned previously, followed by a deep dive into the analysis of treaps and skip lists. Before going into detail, we need to recall important results of the last session:

- The expected time for search for treaps is $O(\log n)$.
- The height of skip lists is $O(\log n)$ with high probability.
- In a skip list, the probability that the key x appears **above** level τ (the lowest level is level 1) is:
$$\Pr[H(x) > \tau] = \frac{1}{2^\tau} \tag{1}$$
It's equivalent to the experiment of repeatedly flipping a fair coin and having the first τ tosses come up heads.
- For any constant $c \geq 1$, we also have the following bound for overall height H of a skip list on n items:
$$\Pr[H > (c+1)\log n] = \Pr[\max_x\{H(x)\} > (c+1)\log n] \leq \frac{1}{n^c}.$$

2 Skip Lists

2.1 Analysis of Randomized Skip List Height

Before showing the expected number of levels in the skip lists, we need to introduce a lemma.

Lemma 1. *For any non-negative integer random variable X:*

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

Before we prove Lemma 1, first, let us prove a simple fact about summations:

Lemma 2.

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

Proof.

$$\begin{aligned}
\sum_{x=1}^{n} x \cdot f(x) &= 1 \cdot f(1) + 2 \cdot f(2) + \dots n \cdot f(n) \qquad \text{Rearranging the terms, we get:} \\
&= \Big(f(1) + f(2) + f(3) + \dots + f(n)\Big) + \Big(f(2) + f(3) + \dots f(n)\Big) + \dots + \Big(f(n)\Big) \\
&= \sum_{k=1}^{n} f(k) + \sum_{k=2}^{n} f(k) + \dots + \sum_{k=n}^{n} f(k) \\
&= \sum_{x=1}^{n} \sum_{k=x}^{n} f(k)
\end{aligned}$$

□

Now we are ready to prove Lemma 1.

Proof of Lemma 1. Recall that $\mathbb{E}[X] = \sum_{x=1}^{\infty} x \cdot \Pr[X = x]$. Then by using the function $f(x) = \Pr[X = x]$ and setting $n = \infty$ in Lemma 2 we get:

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} x \cdot \Pr[X = x] = \sum_{x=1}^{\infty} \sum_{k=x}^{\infty} \Pr[X = k]$$

Now observe that by the definition of probabilities

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

and the lemma follows. □

Based on Lemma 1 and by defining $H = \max_{x}\{H(x)\}$, we have:

$$\begin{aligned}
\mathbb{E}[H] &= \sum_{\tau=1}^{\infty} \Pr[H \geq \tau] \\
&= \sum_{\tau=1}^{\log n} \Pr[H \geq \tau] + \sum_{\tau=\log n+1}^{\infty} \Pr[H \geq \tau]
\end{aligned}$$

Observe that any probability is at most 1, i.e. $\Pr[H \geq \tau] \leq 1$, which will be sufficient for the first term above. To bound the second term, we can prove a tighter bound on $\Pr[H \geq \tau]$ using the Union Bound:

$$
\begin{aligned}
\Pr[H \geq \tau] &= \Pr[H(x_1) \vee \ldots \vee H(x_n)] \\
&\leq \sum_{i=1}^{n} \Pr[H(x_i) \geq \tau)] && \text{By equation (1):} \\
&= \sum_{i=1}^{n} \frac{1}{2^\tau} \\
&= \frac{n}{2^\tau}
\end{aligned}
$$

Then:

$$
\begin{aligned}
\mathbb{E}[H] &= \sum_{\tau=1}^{\log n} \Pr[H \geq \tau] + \sum_{\tau=\log n+1}^{\infty} \Pr[H \geq \tau] \\
&\leq \sum_{\tau=1}^{\log n} 1 + \sum_{\tau=\log n+1}^{\infty} \frac{n}{2^\tau} \\
&= \log n + \sum_{i=1}^{\infty} \frac{n}{2^{\log n+i}} \\
&= \log n + \sum_{i=1}^{\infty} \frac{n}{n \cdot 2^i} \\
&= \log n + 1
\end{aligned}
$$

I.e., the expected number of levels of a skip list is one more than the ideal case where each level contains exactly half the nodes of the level below.

2.2 Search Time

Algorithm 1

function SEARCH(k, S) ▷ Returns the node with the largest key smaller or equal to k
    $v \leftarrow S.start$
    while $(v.next \neq nil$ and $v.next.key \leq k)$ or $(v.down \neq nil)$
        if $v.next \neq nil$ and $v.next.key \leq k$
            $v \leftarrow v.next$
        else
            $v \leftarrow v.down$
    return v

Let us analyze the expected cost of a search in a skip list. This is proportional to the number of elements that the search algorithm will visit. To simplify this computation, consider the sequence of elements that the search algorithm will visit and let us trace this path in the backward direction from the search destination at the bottom of the data structure to the root (see Figure 1 for an illustration). Tracing the algorithm from the search destination node v at the bottom of the level of the data structure is equivalent to the following algorithm:

Algorithm 2 An equivalent algorithm for backward tracing of a search

function BACKWARD-SEARCH(v)
    while $v \neq root$
        if $up(v) \neq NIL$
            $v = up(v)$
        else
            $v = left(v)$

For every node v in the skip list, the link $up(v)$ exists if the coin flip came up HEADS during the insertion of v. Thus, it exists with probability $\frac{1}{2}$. From the analysis perspective, this is exactly like the following algorithm in which we flip a fair coin during the backward traversal.

Algorithm 3 An equivalent algorithm for Backward tracing of a search via coin flipping

function BACKWARD-SEARCH-FLIP(v)
    while $v \neq root$
        if COINFLIP() == HEADS
            $v = up(v)$
        else
            $v = left(v)$

For the above algorithm, we have the same expected number of TAILS and HEADS since we flip a fair coin. Therefore, the expected running time of search, $\mathbb{E}[T]$, is twice the expected number of times of going upward in the tree (the height of the tree L):

$$\mathbb{E}[T] = \mathbb{E}[2L] \leq 2(\log n + 1) = O(\log n)$$

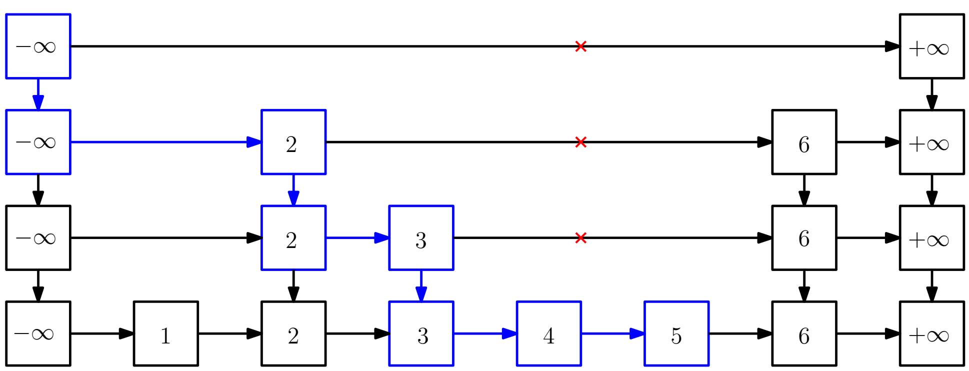

Figure 1: Searching for 5 in a skip list.

3 Treaps

In order to analyze the performance of treaps, we need some elementary techniques for bounding probabilities of random variables deviating from the expected value.

3.1 Markov's Inequality

Markov's inequality provides the most elementary bound which is defined for a non-negative random variables with finite mean as follows:

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

Proof. We can use Lemma 1 to prove this inequality:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{x=0}^{a-1} x \cdot \Pr[X = x] + \sum_{x=a}^{\infty} x \cdot \Pr[X = x] \\
&\geq \sum_{x=a}^{\infty} x \cdot \Pr[X = x] \\
&\geq a \sum_{x=a}^{\infty} \Pr[X = x] \\
&= a \Pr[X \geq a]
\end{aligned}$$

by rearranging the terms:

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

□

3.2 The Height of Treaps

Now, we want to apply Markov's inequality to treaps and see for what values of h, the following inequality holds:

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

The height of a treap is greater than or equal to h the height of at least one node in the treap is greater than or equal to h, i.e.:

$$\Pr[\textsc{Height}(x_1) \geq h \vee \textsc{Height}(x_2) \geq h \vee \ldots \vee \textsc{Height}(x_n) \geq h] \leq \frac{1}{n}$$

The **union bound** states that the probability that at least one of the events happens is no greater than the sum of the probabilities of the individual events. Therefore, we can find an upper bound based on x_k as the node with the kth smallest search key:

$$n \cdot \Pr[\textsc{Height}(x_k) \geq h] \leq \frac{1}{n} \quad \Rightarrow \quad \Pr[\textsc{Height}(x_k) \geq h] \leq \frac{1}{n^2}$$

by applying Markov's inequality:

$$\Pr[\textsc{Height}(x_k) \geq h] \leq \frac{\mathbb{E}[\textsc{Height}(x_k)]}{h}$$

In the previous lecture we showed that $\mathbb{E}[\textsc{Height}(x_k)] \leq 2 \ln n$. Thefore:

$$\frac{\mathbb{E}[\textsc{Height}(x_k)]}{h} \leq \frac{2 \ln n}{h} \leq \frac{1}{n^2}$$

which gives $h \geq 2n^2 \ln n$. As we know, the height of a treap with n elements cannot be greater than n. Thus, the resulting condition for h is not useful. To conclude, the Markov's inequality does not give a tight bound on the value of the height. In the rest, we'll introduce tighter bounds, which provide useful information about the tree structure.

3.3 Chebyshev's inequality

Chebyshev's inequality provides an upper bound to the probability of the absolute deviation between the random variable and its mean. Before we state the inequality, let us define some terminology.

Definition 3. *A set of n random variables $X_1, \ldots, X_n$ are* fully independent *iff*

$$\Pr[(X_1 = x_1) \wedge (X_2 = x_2) \wedge \cdots \wedge (X_n = x_n)] = \Pr[X_1 = x_1] \cdot \Pr[X_2 = x_2] \cdot \cdots \cdot \Pr[X_n = x_n]$$

for all possible values $x_1, x_2, \ldots, x_n$.

Definition 4. *A set of random variables $X_1, X_2, \ldots, X_n$ are k-wise independent if every subset of k variables is independent.*

⟨ Example of pairwise independence vs 3-wise independence missing. ⟩

Corollary 5. *If $X_1, X_2, \ldots, X_n$ are k-wise independent, then for any integer $2 \leq k' < k$, they are also k'-wise independent.*

Theorem 6. *Let $X_1, X_2, \ldots, X_n$ be pairwise independent indicator random variables and $X = \sum_{i=1}^{n} X_i$ with $\mathbb{E}[X] = \mu$, then for any $a > 0$:*

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

When $a = k^2$

$$\Pr[(x - \mu)^2 \geq k^2] < \frac{\mu}{k^2}.$$

Because $(x - \mu)^2 \geq k^2$, we have $|x - \mu| \geq k$ and therefore:

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

By setting $\Delta = k$, we get the *additive* tail bounds:

$$\Pr[x \geq \mu + \Delta] \leq \frac{\mu}{\Delta^2}, \quad \Pr[x \leq \mu - \Delta] \leq \frac{\mu}{\Delta^2}$$

By setting $\delta\mu = k$, we get the *multiplicative* tail bounds:

$$\Pr[x \geq (1 + \delta)\mu] \leq \frac{1}{\delta^2 \mu}, \quad \Pr[x \leq (1 - \delta)\mu] \leq \frac{1}{\delta^2 \mu}.$$

3.3.1 Higher Moment Inequality

Theorem 7. *For any fixed integer $k > 0$, if $X_1, X_2 \ldots X_n$ are $2k$-wise independent and $X = \sum_{i=1}^{n} X_i$, then*

$$\Pr[(X-\mu)^k \geq a] = O\left(\frac{\mu^k}{a}\right).$$

Corollary 8. *The additive bounds for any $\Delta > 0$:*

$$\Pr[X \geq \mu + \Delta] = O\left(\left(\frac{\mu}{\Delta^2}\right)^k\right), \quad \Pr[X \leq \mu - \Delta] = O\left(\left(\frac{\mu}{\Delta^2}\right)^k\right).$$

The multiplicative bounds for any $\delta > 0$:

$$\Pr[X \geq (1+\delta)\mu] = O\left(\left(\frac{1}{\delta^2\mu}\right)^k\right), \quad \Pr[X \leq (1-\delta)\mu] = O\left(\left(\frac{1}{\delta^2\mu}\right)^k\right).$$

3.4 Chernoff bound

If we know that indicator random variables are fully independent, Chernoff bound provides us with even tighter bounds.

Theorem 9. *If indicator random variables $X_1, X_2, \ldots, X_n$ are fully independent, then for all $a \geq \mu$, $X = \sum_{i=1}^{n} X_i$:*

$$\Pr[X \geq a] \leq e^{a-\mu}\left(\frac{\mu}{a}\right)^a \quad \textit{for any } a \geq \mu \textit{ (upper tail)}$$
$$\Pr[X \leq a] \leq e^{a-\mu}\left(\frac{\mu}{a}\right)^a \quad \textit{for any } a \leq \mu \textit{ (lower tail)}$$

Corollary 10. *The additive tail bounds for any $\Delta > 0$:*

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

The multiplicative tail bounds for any $\delta > 0$:

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

If we know that $0 < \delta < 1$, then we can derive a simpler bound:

Corollary 11. *The Chernoff multiplicative bounds for any $0 < \delta < 1$:*

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

This simple form of Chernoff bound is easier to use, but we should notice that they provide probability with looser bounds.