Let $S = \{x_1 = (k_1, p_1), \ldots, x_n = (k_n, p_n)\}$ be a set of items (keys and priorities) used to construct a max-treap $T(S)$. The $p_i$ are independent $U[0,1]$ random variates, and we assume the keys are unique and sort the items in decreasing order of keys, writing $x_1 > \ldots > x_n$.

1. Outline

To bound the expected height of a treap, we will show a high probability bound on the depth of any element in the treap, use a union bound to convert this into a high probability bound on the maximum depth of any element in the treap, then argue this implies the expected height is small.

2. High probability bound on height of a treap

For the first step of establishing a high probability bound on the depth of an element $x_i$ in the treap, we recall that
\begin{equation}
\text{depth}(x_i) = \sum_{j=1}^{i-1} \xi_j + \sum_{j=i+1}^{n} \xi_j,
\end{equation}
where $\xi_j = \mathbb{1}_{x_j}$ is an ancestor of $x_i$. Recall that $x_j$ is an ancestor of $x_i$ if and only if $p_j$ is the largest priority in the ordered set of items $(x_j, \ldots, x_i)$. Thus knowledge of $\xi_j$ is equivalent to knowing whether or not $p_j$ is the largest priority in the ordered set of items $(x_j, \ldots, x_i)$.

Note that
\[
\mathbb{P}\left[\bigcap_{j=1}^{i-1} \{\xi_j = \nu_j\}\right] = \prod_{j=1}^{i-1} \mathbb{P}[\xi_j = \nu_j | \xi_{j+1} = \nu_{j+1}, \ldots, \xi_{i-1} = \nu_{i-1}] = \prod_{j=1}^{i-1} \mathbb{P}[\xi_j = \nu_j],
\]
because knowing the index of the largest priorities in the subset $(p_{j+1}, \ldots, p_{i-1})$ gives no information about whether or not $p_j$ is larger than $\max(p_{j+1}, \ldots, p_{i-1})$. Thus the Poisson trials in the first sum on the right-hand side of (1) are independent. A similar argument shows the Poisson trials in the second sum are independent.

Fix a constant $C > 6$. If the depth of an element is $d$, then at least one of the sums on the righthand-side of (1) must be greater than $d/2$, so
\[
\mathbb{P}[\text{depth}(x_i) \geq 2C\mathbb{E}[\text{depth}(x_i)]] \leq \mathbb{P}\left[\sum_{j=1}^{i-1} \xi_j \geq C\mathbb{E}[\text{depth}(x_i)]\right] + \mathbb{P}\left[\sum_{j=i+1}^{n} \xi_j \geq C\mathbb{E}[\text{depth}(x_i)]\right] \leq 2^{1-C\mathbb{E}[\text{depth}(x_i)]}
\]
where the final inequality follows from a Chernoff bound for Poisson trials and the fact that the expectations of the sums of Poisson trials are both smaller than the expected depth. Now we use the fact\(^2\) that $H_n \geq \ln(n+1)$ to bound the expected depth of $x_i$ below:
\[
\mathbb{E}[\text{depth}(x_i)] = H_{n-i+1} + H_i - 2 \\
\geq \ln(n - i + 2) + \ln(i + 1) \\
\geq \max(\ln(n - i), \ln(i + 1)) \\
\geq \ln\left(\frac{n}{2}\right) \\
= \ln(2)(\log_2(n) - 1)
\]
\(^1\)However, the two sets of summands are not independent of each other. One way to see this is to note that if the first sum is zero, then we know that $p_i$ is the maximum of $i$ independent random $U[0,1]$ variates, so conditioning on that, when $j > i$, by asking if $p_j$ is the largest of $(p_{i+1}, \ldots, p_j)$, we are essentially asking that $p_j$ not just be the maximum of $j - i + 1$ independent copies of $U[0,1]$, but that it is the maximum of $j$ independent copies. This example shows that, in general, the distribution of $\xi_j$ conditioned on knowledge of the first sum changes from the unconditioned distribution.
\(^2\)This can be shown by comparing $H_n$ to an integral.
and recall our previous estimate

$$E[\text{depth}(x_i)] \leq 2 \ln(n).$$

Using these two estimates in (2) gives

$$\mathbb{P}[\text{depth}(x_i) \geq 4C \ln(n)] \leq \mathbb{P}[\text{depth}(x_i) \geq 2C E[\text{depth}(x_i)]]$$

$$\leq 2^{1-C E[\text{depth}(x_i)]}$$

$$\leq 2^{1+C \ln(2)-C \ln(2) \log_2(n)}$$

$$\leq 2^{1+C \frac{1}{n^{C/2}}}$$

Now we can take $C = 6$ and conclude that

$$\mathbb{P}[\text{depth}(x_i) \geq 24 \ln(n)] \leq \frac{128}{n^3},$$

which is meaningful when $n \geq 6$. These constants are not optimal— you can be more careful and get tighter bounds— but we only care about asymptotics here, so this is good enough.

A union bound shows

$$\mathbb{P}[\max_{i=1,\ldots,n} \text{depth}(x_i) \geq 24 \ln(n)] \leq \frac{128}{n^2},$$

which is meaningful when $n \geq 12$.

3. Expected height of a treap

The fact that the expected height of a treap is $O(\ln(n))$ follows immediately from (4). Let $H = \max_{i=1,\ldots,n} \text{depth}(x_i)$ be the height of the treap, then

$$E[H] = \sum_{\ell=1}^{24 \ln(n)} \ell \mathbb{P}[H = \ell] + \sum_{\ell=25 \ln(n)+1}^{n-1} \ell \mathbb{P}[H = \ell]$$

$$\leq 24 \ln(n) \mathbb{P}[H \leq 24 \ln(n)] + (n-1) \mathbb{P}[H \geq 24 \ln(n)]$$

$$\leq 24 \ln(n) + \frac{128}{n} = O(\ln(n)).$$