Exercise sheet 11.

Data structures SoSe 2020

László Kozma, Katharina Klost

Due 10:00, July 13th, 2020

Exercise 1 GreedyFuture *4+4 Points*

Recall the GreedyFuture BST algorithm and its geometric version (called simply Greedy).

(a) Show that the Greedy algorithm achieves $O(n)$ cost on the search sequence $R = 1, \ldots, n$. Show this in geometric view, then verify what happens in tree view. In geometric view you can take as cost the total number of points in the point set obtained in the end.

(b) *Bonus (+4p)*: Extend the $O(n)$ bound of the previous question to sequences R that are "V-type". A V-type sequence is such that every entry is either larger than all previous entries or smaller than all previous entries. For example, $R = 5, 4, 3, 6, 2, 1, 7, 8, 9$ is a V-type sequence. Can you show such a bound if R is the reverse of a V-type sequence?

Both questions can be answered in geometric view, but it is useful to check what happens in tree-view as well.

(c) In this part, we want to show that the cost of Greedy can be more than the optimum. Consider the search sequence $R = 2, 1, 3, 1, 3, 1, 3, 1, 3, \ldots$. What is the cost of Greedy when serving R? What is $OPT(R)$? You can answer both questions in the geometric view. How do the two executions differ in tree-view?

Exercise 2 A difficult sequence *1+3+1+3 Points*

In the previous exercise sheet we showed that for most sequences R of size n, we have $OPT(R) \in \Theta(n \log n)$, i.e. most sequences are "difficult" for all BSTs. Now we look at a *concrete* hard sequence.

Let $R_n = r_0, \ldots, r_{n-1}$ be a permutation of the integers $\{0, \ldots, n-1\}$, where $n = 2^k$ for some k.

Each entry r_i is obtained by reversing the binary representation of the value i, padded with zeros to exactly k bits. For example, in R_{16} we have $r_{11} = 13$ and $r_1 = 8$, since 11 and 13 are 1011 and 1101 in binary, and 1 and 8 are 0001 and 1000.

(a) Write the sequence R_{16} and plot it in geometric view.

(b) Consider R_8 as a search sequence, and plot the geometric view of executing R_8 by: (i) a fixed static tree of your choice, (ii) Greedy, (iii) Move-to-root or Splay (whichever you choose) with an arbitrary initial tree.

What are the costs of these executions? (By cost we mean the total number of points in geometric view.)

(c) Compute the interleave lower bound $IL(R_8)$.

(d) Show that $OPT(R_n) \in \Theta(n \log n)$.

Exercise 3 Interleave lower bound *4 Points*

Let $IL(R)$ denote the interleave lower bound for a sequence R. Our goal is to show that in general, this lower bound is not tight.

Show that there is an arbitrarily long sequence R over n keys, for which $OPT(R)/IL(R) = \Omega(\log \log n)$.

Total: 20 points. Have fun with the solutions!