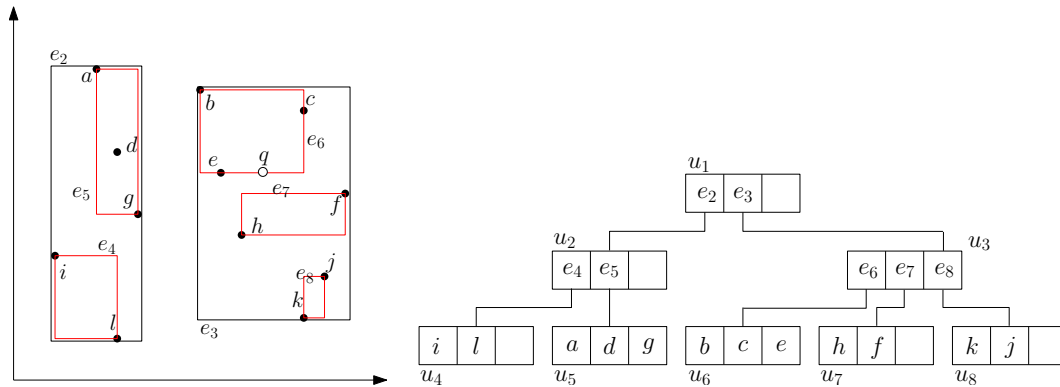


## WST540: Exercise 6

**Problem 1.** Consider the following R-tree:



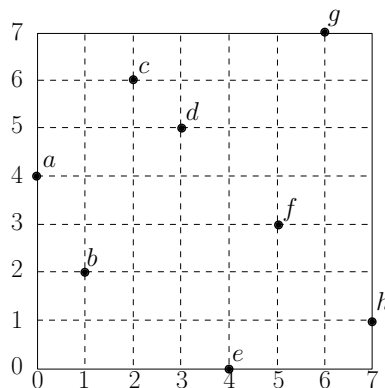
Suppose that we apply the best-first algorithm to find the nearest neighbor of the query point  $q$  as shown in the picture. List the nodes in the order that they are visited by the algorithm.

**Solution.**  $u_1, u_3, u_6, u_7$ . The algorithm then terminates by reporting  $e$ .

**Problem 2.** Repeat the above by finding the 2 nearest neighbors of  $q$ .

**Solution.**  $u_1, u_3, u_6, u_7$ . The algorithm reports  $e$ , then  $h$ , and then terminates.

**Problem 3.** Calculate the z-values of the black points in the following figure (the data space has domain  $[0, 7]$  on each dimension):

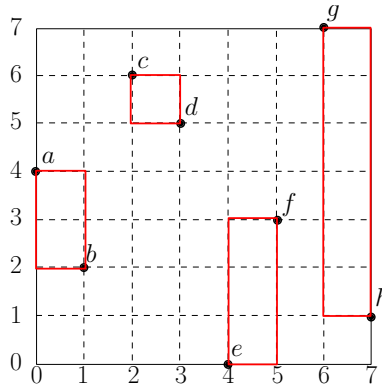


**Solution.**

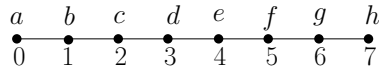
point	z-value
$a$	$(010000)_2 = 16$
$b$	$(000110)_2 = 6$
$c$	$(011100)_2 = 28$
$d$	$(011011)_2 = 27$
$e$	$(100000)_2 = 32$
$f$	$(100111)_2 = 39$
$g$	$(111101)_2 = 61$
$h$	$(101011)_2 = 43$

**Problem 4.** Consider that we create an R-tree on the points in the previous problem using the method discussed in our lecture. Show the leaf MBRs of the R-tree.

**Solution.**



**Problem 5.** Consider that a server hosts a 1d hidden dataset  $D$  which contains 8 points as shown below. We want to discover the entire  $D$  by issuing range queries in the way described in class. Suppose that the value of  $k$  is 4, such that whenever the query result has more than 4 points, the server always returns the first 4 points alphabetically (e.g., for a query with range  $[2, 7]$ , the server returns  $c, d, e, f$ ). Give the queries that need to be issued by our algorithm.



**Solution.** We start by issuing  $q_1 = (-\infty, \infty)$ . The server returns  $\{a, b, c, d\}$ . Hence,  $q_1$  is divided by a 3-way split into  $q_2 = (-\infty, 0]$ ,  $q_3 = [1, 1]$ , and  $q_4 = [2, \infty)$ . Queries with  $q_2$  and  $q_3$  are resolved. For  $q_4$ , the server returns  $\{c, d, e, f\}$ . Hence,  $q_4$  is divided into  $q_5 = [2, 2]$ ,  $q_6 = [3, 3]$ , and  $q_7 = [4, \infty)$ . The queries with these three intervals are all resolved.