

Programming Assignment (optional)

due: January 11th, 2018, 23:59

Kd-trees are often used for answering queries other than orthogonal range search. One of the most common usages of kd-trees is for nearest-neighbor search: We preprocess a finite set P of points into a d -dimensional kd-tree and then answer queries of the form “given a query point $q \in \mathbb{R}^d$, find the k points of P that are closest, in Euclidean distance, to q among all points of P .”

The project is to implement the construction of a kd-tree for that purpose. The input consists of two integers $2 \leq d \leq 15$ and $n > 0$, followed by n d -tuples of rational numbers (pairs of integers), representing the set P of points. A query is then a pair q and k , where q is a point in \mathbb{R}^d given as a d -tuple of rational coordinates, and $k > 0$ is the number of closest points in P that we wish to find. The algorithm should always return at most k points: in case of a tie, the lexicographically k smallest points should be returned.

Within the kd-tree framework you still have considerable freedom. For example, you can choose a different splitting criterion, other than the median in the specific coordinate. You can alternate differently among the different coordinates and not necessarily in the standard order d_1, d_2, \dots . You may choose to store the data points in nodes other than the leaves. However, you need to justify in an accompanying text every non-basic choice that you make (by basic we mean the procedure as it is described in the textbook).

You are required to implement the construction of the data structure and the query algorithm from scratch. You may use existing libraries for any primitive operation that you need. You are encouraged to write the code in C++ and use CGAL. If you wish to write in another language you should first get the approval of the course grader Shahar Shamaï.

More details on the input/output will be provided in the course website. Notice in particular that in CGAL a rational number given as a pair of integers has a separating slash between them; for details see the manual.

Your implementation will be benchmarked for time efficiency (both the construction and the query) and the results will be published in the course website.

There will be a competition for efficiency among all the submissions.