COP 5405: Advanced Algorithms

Fall 2006

Lecture 19

Higher dimensional range trees

A *d*-dimensional range tree can be constructed as follows, for $d \geq 2$.

- The main tree is a balanced BST on the x_1 coordinates.
- The associated data structure for node *v* in the above tree is a *d-1* dimensional range tree on the $(x_2, ..., x_d)$ coordinates.
    - The associated data structure stores the canonical subset of points *P(v)*.
- Construction time is $O(n \log n)$ (review question), and storage too is $O(n \log n)$.

Querying a *d*-dimensional range tree is similar to that of a *2*-dimensional tree, except that we replace the call to a *1*-dimensional range query with that to a *d-1* dimensional range query.

- Query time is $O(\log^d n + k)$ to report *k* points, $d \geq 2$.

The complexity of the overhead for querying (excluding the time for reporting points) can be proven using induction, based on the following two observations:

- $Q_d(n) = O(\log n) + O(\log n)\, Q_{d-1}(n),\ d \geq 2$
- $Q_2(n) = \log^2 n$

The cost of reporting *k* points is $O(k)$.

General set of points

We wish to remove the assumption that no two points have identical values for the *x* and *y* coordinates. We accomplish this by replacing real numbers by composite numbers as follows.

- A composite number has the form $(a|b)$, where *a* and *b* are reals.
- We impose a total order on the composite numbers by defining $(a|b) < (a'|b')$ iff $a < a'$ or ($a=a'$ and $b < b'$).
- We replace each point (x, y) by $(x|y, y|x)$.
    - Now, each point has distinct *x* and *y* coordinates.
- A query range $[x, x'] \times [y, y']$ is transformed to the query range $[x|-\infty, x'|\infty] \times [y|-\infty, y'|\infty]$.