

# More Branch and Bound Algorithms

Algorithmic Problems Around the Web #3

Yury Lifshits

<http://yury.name>

CalTech, Fall'07, CS101.2, <http://yury.name/algoweb.html>

1 / 19

## Outline

- 1 Variations of Metric Trees
- 2 M-Trees
- 3 Branch and Bound for Euclidean Space

2 / 19

## Part I

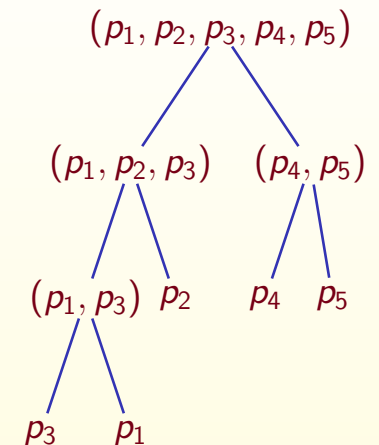
### Variations of Metric Trees

3 / 19

## Branch and Bound: Range Search

**Task:** find all  $i$   $d(p_i, q) \leq r$ :

- 1 Make a depth-first traversal of search hierarchy
- 2 At every node compute the lower bound for its subtree
- 3 Prune branches with lower bounds above  $r$

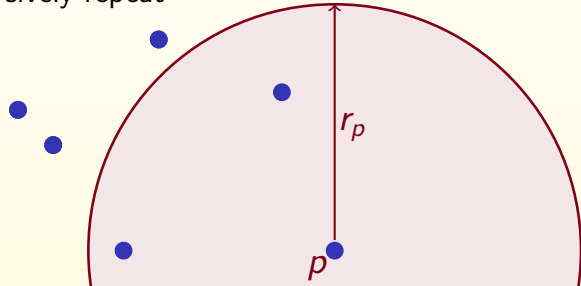


4 / 19

# Vantage-Point Partitioning

Uhlmann'91, Yianilos'93:

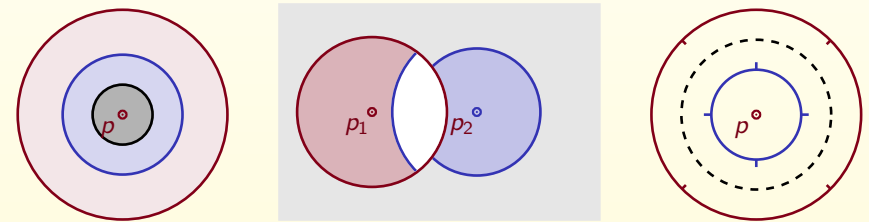
- 1 Choose some object  $p$  in database (called **pivot**)
- 2 Choose partitioning radius  $r_p$
- 3 Put all  $p_i$  such that  $d(p_i, p) \leq r$  into "inner" part, others to the "outer" part
- 4 Recursively repeat



5 / 19

# Variations of Vantage-Point Trees

- **Burkhard-Keller tree:** pivot used to divide the space into  $m$  rings Burkhard&Keller'73
- **MVP-tree:** use the same pivot for different nodes in one level Bozkaya&Ozsoyoglu'97
- **Post-office tree:** use  $r_p + \delta$  for inner branch,  $r_p - \delta$  for outer branch McNutt'72

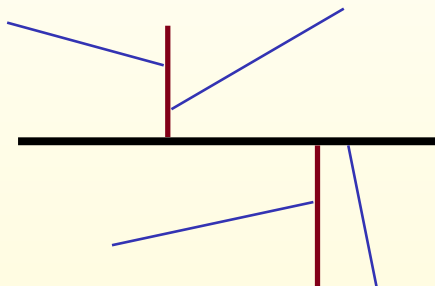


6 / 19

# Generalized Hyperplane Tree

Partitioning technique (Uhlmann'91):

- Pick two objects (called pivots)  $p_1$  and  $p_2$
- Put all objects that are closer to  $p_1$  than to  $p_2$  to the left branch, others to the right branch
- Recursively repeat



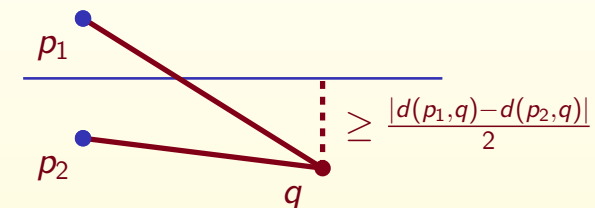
7 / 19

# GH-Tree: Pruning Conditions

**For  $r$ -range search:**

- If  $d(q, p_1) > d(q, p_2) + 2r$  prune the left branch
- If  $d(q, p_1) < d(q, p_2) - 2r$  prune the right branch

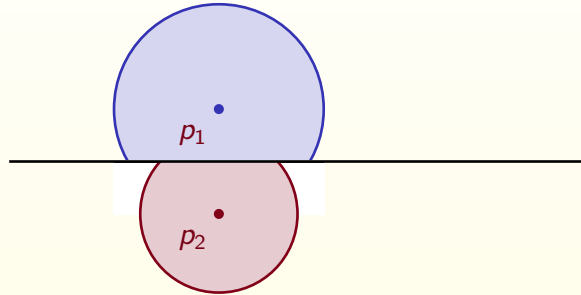
For  $|d(q, p_1) - d(q, p_2)| \leq 2r$  we have to inspect both branches



8 / 19

## Bisector trees

Let's keep the covering radius for  $p_1$  and left branch, for  $p_2$  and right branch: useful information for stronger pruning conditions



**Variation:** monotonous bisector tree (Noltemeier, Verbag, Zirkelbach'92) always uses parent pivot as one of two children pivots

**Exercise:** prove that covering radii are monotonically decrease in mb-trees

9 / 19

## Part II

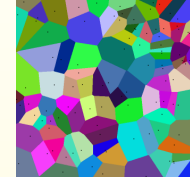
## M-trees

11 / 19

## Geometric Near-Neighbor Access Tree

Brin'95:

- Use  $m$  pivots
- Branch  $i$  consists of objects for which  $p_i$  is the closest pivot
- Stores minimal and maximal distances from pivots to all "brother"-branches



10 / 19

## M-tree: Data structure

Ciaccia, Patella, Zezula'97:

- All database objects are stored in leaf nodes (buckets of fixed size)
- Every internal nodes has associated pivot, covering radius and legal range for number of children (e.g. 2-3)
- Usual depth-first or best-first search

Special algorithms for insertions and deletions a-la B-tree

12 / 19

## M-tree: Insertions

All insertions happen at the leaf nodes:

- 1 Choose the leaf node using “minimal expansion of covering radius” principle
- 2 If the leaf node contains fewer than the maximum legal number of elements, there is room for one more. Insert; update all covering radii
- 3 Otherwise the leaf node is split into two nodes
  - 1 Use two pivots generalized hyperplane partitioning
  - 2 Both pivots are added to the node’s parent, which may cause it to be split, and so on

13 / 19

## Part III

### k-d Trees, R-trees

14 / 19

## Advantages of Euclidean Space

- Rich mathematical formalisms for defining a boundary of any set  
**Examples:** rectangles, hyperplanes, polynomial curves
- Easy computation of lower bound on distance between query point and any set boundary
- Easy definable mappings to smaller spaces

15 / 19

## k-d Tree

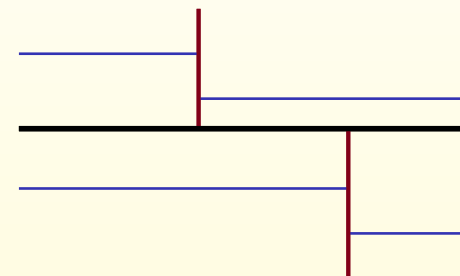
### Preprocessing:

Bentley, 1975

Top-down partitioning  
On level  $l$ : split the current set by hyperplane orthogonal to  $l \bmod k$  axis

### Query processing:

Standard branch and bound



16 / 19

# R-Tree

## Preprocessing:

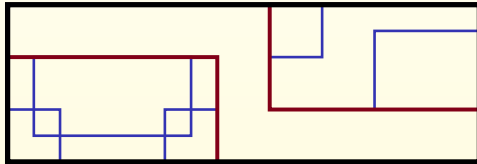
Guttman, 1984

- Bottom-up partitioning
- Keep bounding rectangles
- Every time: merge current rectangles and compute bounding rectangle for every group

## Query processing:

- Standard branch and bound

**Insertions/deletions:** similar to M-tree, B-tree



17 / 19

Thanks for your attention! Questions?


18 / 19

# References

Course homepage <http://yury.name/algoweb.html>

 P. Zezula, G. Amato, V. Dohnal, M. Batko  
Similarity Search: The Metric Space Approach. Springer, 2006.  
<http://www.nmis.isti.cnr.it/amato/similarity-search-book/>

 E. Chávez, G. Navarro, R. Baeza-Yates, J. L. Marroquín  
Searching in Metric Spaces. ACM Computing Surveys, 2001.  
<http://www.cs.ust.hk/~leichen/courses/comp630j/readings/acm-survey/searchinmetric.pdf>

 G.R. Hjaltason, H. Samet  
Index-driven similarity search in metric spaces. ACM Transactions on Database Systems, 2003  
[http://www.cs.utexas.edu/~abhinay/ee382v/Project/Papers/ft\\_gateway.cfm.pdf](http://www.cs.utexas.edu/~abhinay/ee382v/Project/Papers/ft_gateway.cfm.pdf)

19 / 19