PRINCIPLES OF DATA ORGANISATION

R-Tree Greene

MOTIVATION

- How to search effectively in more than one dimension?
- B-tree for multidimensional data ~ R-tree

R-TREE GREENE

- Greene 1989
- Modification of the split algorithm of the original R-tree (Guttman)
- Splitting is based on a hyperplane which defines in which node the objects will fall

INSERT : SPLITNODE

```
SplitNode(P,PP,E)
  ChooseAxis();     // choose the hyperplane
  Distribute();
```

ChooseAxis()

PickSeads; { from Guttman's version – returns seeds E_i and E_j }

For every axis compute the distance between MBRs E_i, E_j;

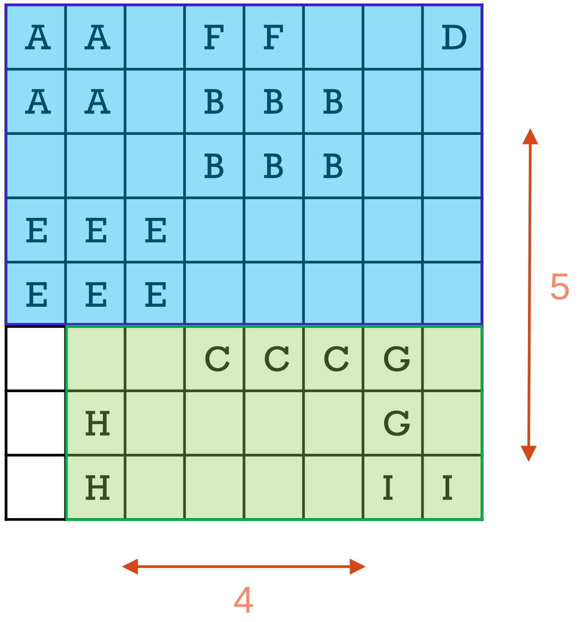
Normalize the distances by the respective edge length of the bounding rectangle of the original node;

Pick the axis with greatest normalized separation;

Distribute()

Sort E_is in the chosen axis j based on the j-th coordinate;

Add first $\lceil (M+1)/2 \rceil$ records into P and rest of them into PP;



PickSeeds:
A and I (Guttman): 64 – 6 = 58

ChooseAxis:
Axis X: 4/8
Axis Y: 5/8 … maximum
Normalization: we have 8x8, but in case of 8x32 the distance should be relative to this size

Distribute:
I, H, G, C and E, B, A, F, D
Note: Or E can bee moved to the other group