Geometric Algorithms

Range searching

Quadtrees, 2D trees, kD trees

Intersections of geometric objects

Geometric search: overview

Types of data: points, lines, planes, polygons, circles, ...
This lecture: sets of N objects.

Geometric problems extend to higher dimensions.
- Good algorithms also extend to higher dimensions.

Basic problems.
- Range searching.
- Nearest neighbor.
- Finding intersections of geometric objects.

1D Range Search

Extension to symbol-table ADT with comparable keys.
- Insert key-value pair.
- Search for key k.
- How many records have keys between k_1 and k_2?
- Iterate over all records with keys between k_1 and k_2.

Application: database queries.

insert B	B
insert D	B D
insert A	A B D
insert I	A B D I
insert H	A B D H I
insert F	A B D F H I
insert P	A B D F H I P
count G to K	2
search G to K	H I

Geometric intuition.
- Keys are point on the line.
- How many points in a given interval?

1D Range Search Implementations

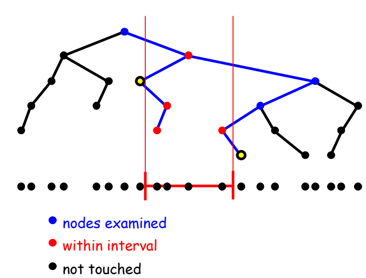
Range search: how many records have keys between k_1 and k_2?

Ordered array. Slow insert, binary search for k_1 and k_2 to find range.
Hash table. No reasonable algorithm (key order lost in hash).

BST. In each node x, maintain number of nodes in tree rooted at x.
Search for smallest element $\geq k_1$ and largest element $\leq k_2$.

	insert	count	range
ordered array	N	log N	R + log N
hash table	1	N	N
BST	log N	log N	R + log N

N = # records
R = # records that match



2D Range Search

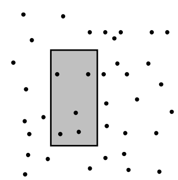
Extension to symbol-table ADT with 2D keys.

- Insert a 2D key.
- Search for a 2D key.
- Range search: find all keys that lie in a 2D range?

Applications: networking, circuit design, databases.

Geometric interpretation.

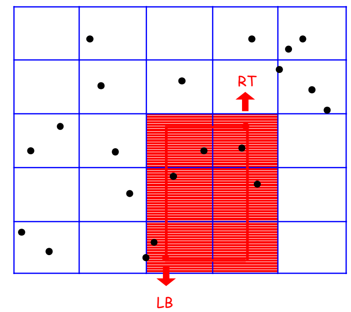
- Keys are point in the plane.
- Find all points in a given h-v rectangle?



2D Range Search Grid Implementation

Grid implementation. (Sedgewick 3.18)

- Divide space into M-by-M grid of squares.
- Create linked list for each square.
- Use 2D array to directly access relevant square.
- Insert: insert (x, y) into corresponding grid square.
- Range search: examine only those grid squares that could have points in the rectangle.



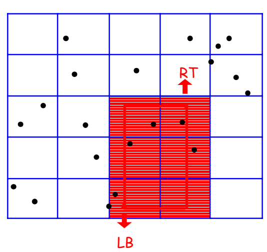
2D Range Search Grid Implementation Costs

Space-time tradeoff.

- Space: $M^2 + N$.
- Time: $1 + N / M^2$ per grid cell examined on average.

Choose grid square size to tune performance.

- Too small: wastes space.
- Too large: too many points per grid square.
- rule of thumb: $\sqrt{N}$ by $\sqrt{N}$ grid.



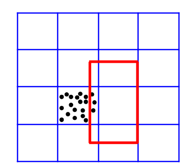
Time costs.

- Initialize: O(N) to initialize 2D array of lists.
- Insert: O(1).
- Range: O(1) per point in range.

⇐ assumes points are evenly distributed

Clustering

Grid implementation. Fast, simple solution for well-distributed points.
Problem. Clustering is a well-known phenomenon in geometric data.



Ex: USA map data.

- 80,000 points, 20,000 grid squares.
- Half the grid squares are empty.
- Half the points have ≥ 10 others in same grid square.
- Ten percent have ≥ 100 others in same grid square.

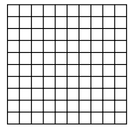
Need data structure that gracefully adapts to data.

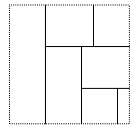
Space Partitioning Trees

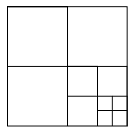
Space partitioning tree. Use a tree to represent the recursive hierarchical subdivision of d-dimensional space.

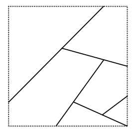
BSP tree. Recursively divide space into two regions.
Quadtree. Recursively divide plane into four quadrants.
Octree. Recursively divide 3D space into eight octants.
kD tree. Recursively divide k-dimensional space into two half-spaces.

Applications.
- Ray tracing.
- Flight simulators.
- N-body simulation.
- Collision detection.
- Astronomical databases.
- Adaptive mesh generation.
- Accelerate rendering in Doom.
- Hidden surface removal and shadow casting.


Grid


kD tree


Quadtree

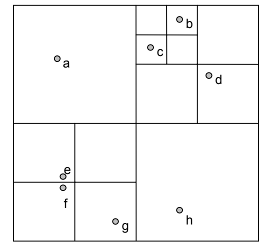

BSP tree

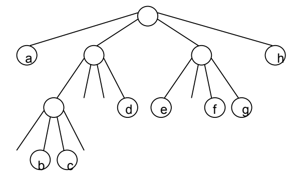
Quad Trees

Quad tree. Recursively partition plane into 4 quadrants.

Implementation: 4-way tree.

```
public class Quadtree {
   Quad quad;
   Object value;
   Quadtree NW, NE, SW, SE;
}
```



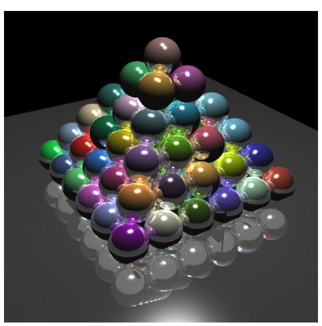


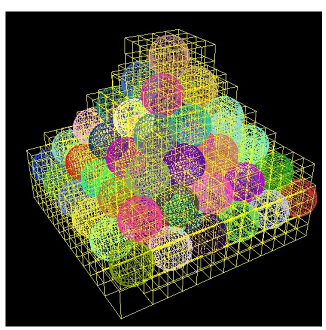
Good clustering performance is a primary reason to choose quad trees over grid methods.

Curse of Dimensionality

Range search / nearest neighbor in k dimensions?
Main application. Multi-dimensional databases.

3D space. Octrees: recursively divide 3D space into 8 octants.
100D space. Centrees: recursively divide into 100 centrants???



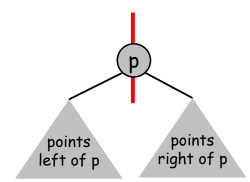


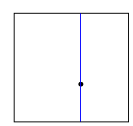
Raytracing with octrees
http://graphics.cs.ucdavis.edu/~gregorsk/graphics/275.html

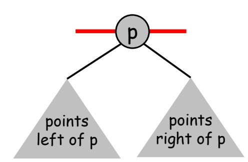
2D Trees

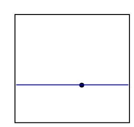
2D tree. Recursively partition plane into 2 halfplanes.

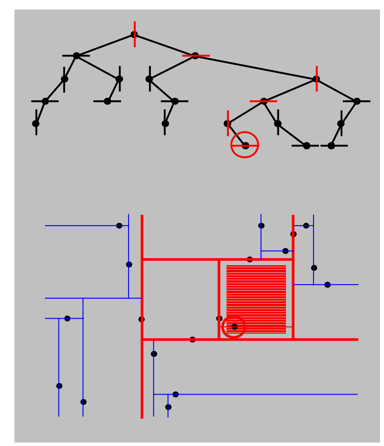
Implementation: BST, but alternate using x and y coordinates as key.
- Search gives rectangle containing point.
- Insert further subdivides the plane.




even levels



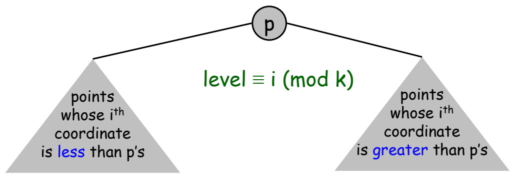

odd levels

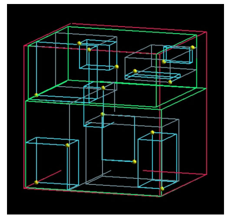


kD Trees

kD tree. Recursively partition k-dimensional space into 2 halfspaces.

Implementation: BST, but cycle through dimensions ala 2D trees.





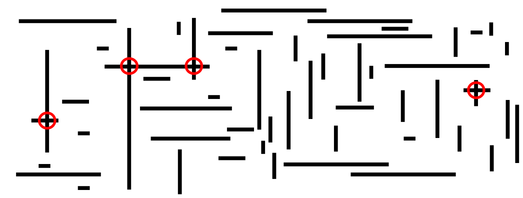
Efficient, simple data structure for processing k-dimensional data.

- Adapts well to high dimensional data.
- Adapts well to clustered data.
- Discovered by an undergrad in an algorithms class!

Geometric Intersection

Problem: find all intersecting pairs among set of N geometric objects.
Applications: CAD, games, movies, virtual reality.

Simple version: 2D, all objects are horizontal or vertical line segments.

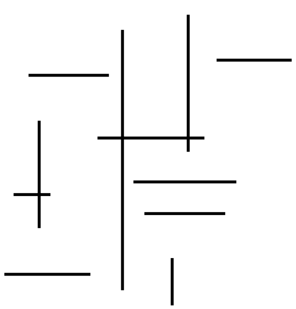


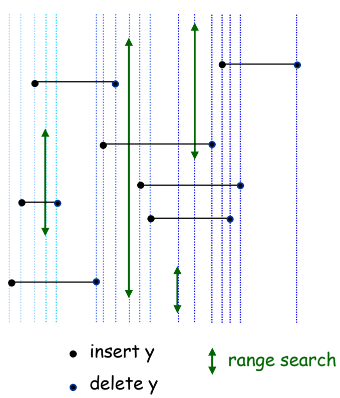
Brute force: test all $\Theta(N^2)$ pairs of line segments for intersection.
Sweep line: efficient solution extends to 3D and general objects.

Orthogonal Segment Intersection: Sweep Line Algorithm

Use horizontal sweep line moving from left to right.

- Sweep line: sort segments by x-coordinate and process in this order.
- Left endpoint of h-segment: insert y coordinate into ST.
- Right endpoint of h-segment: remove y coordinate from ST.
- v-segment: range search for interval of y endpoints.





Orthogonal Segment Intersection: Sweep Line Algorithm

Sweep line reduces 2D orthogonal segment intersection problem to 1D range searching!

Running time of sweep line algorithm.

- Sort by x-coordinate. O(N log N)
- Insert y-coordinate into ST. O(N log N)
- Delete y-coordinate from ST. O(N log N)
- Range search. O(R + N log N)

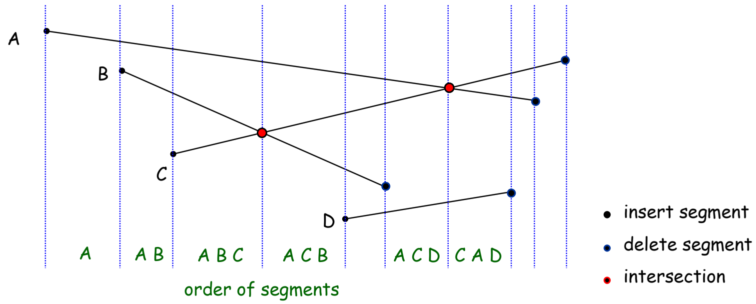
N = # line segments
R = # intersections

Efficiency relies on judicious use of data structures.

Line Segment Intersection: General Version

Use horizontal sweep line moving from left to right.

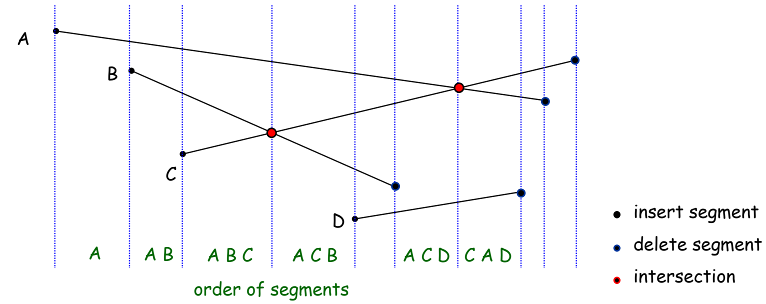
- Maintain order of segments that intersect sweep line by y-coordinate.
- Intersections can only occur between adjacent segments.
- Add/delete line segment $\Rightarrow$ one new pair of adjacent segments.
- Intersection $\Rightarrow$ two new pairs of adjacent segments.



Line Segment Intersection: General Version

Efficient implementation of sweep line algorithm.

- Maintain PQ of important x-coordinates – endpoints and intersections.
- Maintain ST of segments intersecting sweep line, sorted by y.
- O(R log N + N log N).

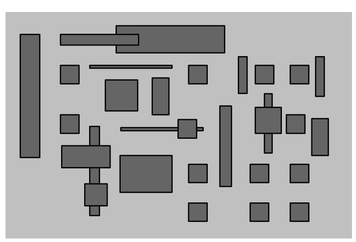


Algorithms and Moore's Law

VLSI database problem: Find all intersections among h-v rectangles.
Application: microprocessor design.

Early 1970s: microprocessor design became a geometric problem

- Very Large Scale Integration (VLSI).
- Computer-Aided Design (CAD).
- Design-rule checking.



Algorithms and Moore's Law

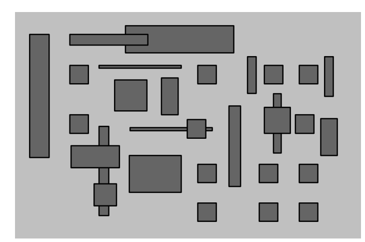
Moore's Law: processing power doubles every 18 months.

- 197x: need to check N rectangles.
- 197(x+1.5): need to check 2N rectangles on a 2x-faster computer.

Quadratic algorithm: compare each rectangle against all others.

- 197x: takes M days.
- 197(x+1.5): takes (4M)/2 = 2M days. (!!)

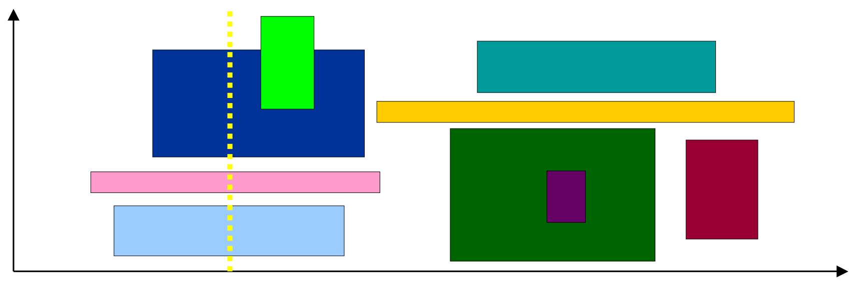
Need O(N log N) CAD algorithms to sustain Moore's Law.



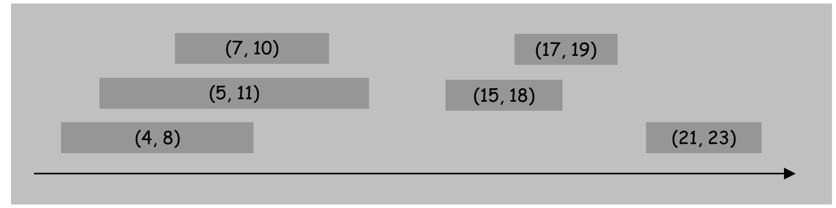
VLSI Database Problem

Move a vertical "sweep line" from left to right.

- Sweep line: sort rectangles by x-coordinate and process in this order.
- Maintain data structure of intervals intersecting sweep line.
- Key operation: given a new interval, does it intersect an existing one?



Interval Search Trees

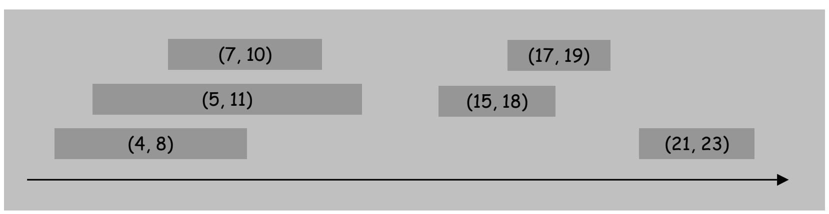


Support following operations.

- Insert an interval `(lo, hi)`.
- Delete the interval `(lo, hi)`.
- Search for an interval that overlaps `(lo, hi)`.

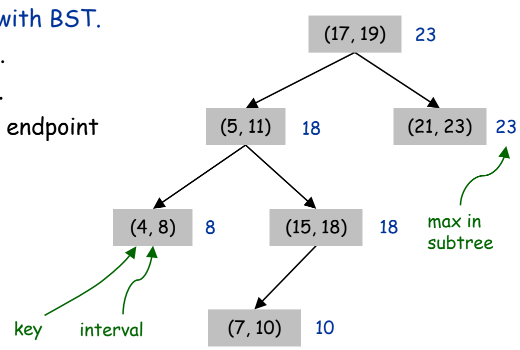
Non-degeneracy assumption. No rectangles share same x-coordinate.

Interval Search Trees



Interval tree implementation with BST.

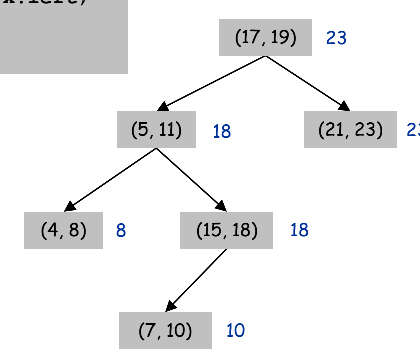
- BST nodes contain interval.
- BST sorted on `lo` endpoint.
- Additional info: store max endpoint in subtree rooted at node.



Finding an Overlapping Interval

Search for an interval that overlaps I = `(lo, hi)`.

```
x = root;
while (x != null) {
   if (x.interval.overlaps(lo, hi))
      return x.interval;
   if      (x.left == null)  x = x.right;
   else if (x.left.max < lo) x = x.right;
   else                      x = x.left;
}
return null;
```



Finding an Overlapping Interval

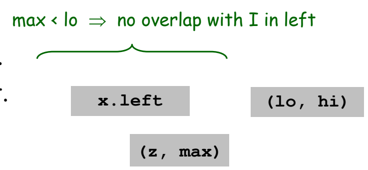
Search for an interval that overlaps I = (lo, hi).

```
x = root;
while (x != null) {
   if (x.interval.overlaps(lo, hi)) return x.interval;
   if      (x.left == null)   x = x.right;
   else if (x.left.max < lo) x = x.right;
   else                       x = x.left;
}
return null;
```

Case 1 (right). If search goes right, then there exists an overlap in right subtree or no overlap in either.

Proof. Suppose no overlap in right.

- (x.left == null) ⇒ no overlap in left.
- (x.left.max < lo) ⇒ no overlap in left.



Finding an Overlapping Interval

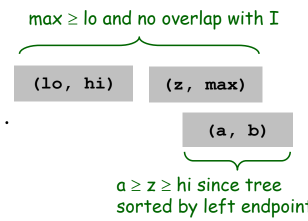
Search for an interval that overlaps I = (lo, hi).

```
x = root;
while (x != null) {
   if (x.interval.overlaps(lo, hi)) return x.interval;
   if      (x.left == null)   x = x.right;
   else if (x.left.max < lo) x = x.right;
   else                       x = x.left;
}
return null;
```

Case 2 (left). If search goes left, then there exists an overlap in left subtree or no overlap in either.

Proof. Suppose no overlap in left.

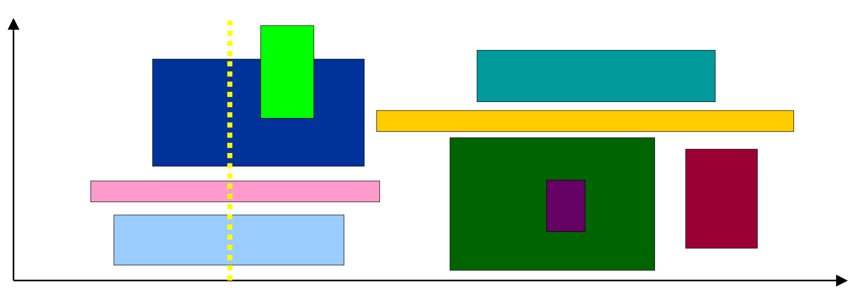
- (x.left.max >= lo) ⇒ no interval (a, b) in right subtree overlaps (lo, hi).



VLSI Database Problem

Move a vertical "sweep line" from left to right.

- Sweep line: sort rectangles by x-coordinates and process in this order, stopping on left and right endpoints.
- Store set of rectangles that intersect the sweep line in an interval search tree (using y-interval of rectangle).
- Left side: interval search for y-interval of rectangle, insert y-interval.
- Right side: delete y-interval.



VLSI Database Problem: Sweep Line Algorithm

Sweep line reduces 2D orthogonal rectangle intersection problem to 1D interval searching!

Running time of sweep line algorithm.

- Sort by x-coordinate. O(N log N)
- Insert y-interval into ST. O(N log N)
- Delete y-interval from ST. O(N log N)
- Interval search. O(R + N log N)

N = # line segments
R = # intersections

Efficiency relies on judicious extension of BST.

Summary

Basis of many geometric algorithms: search in a planar subdivision.

	grid	2D tree	Voronoi diagram	intersecting lines
basis	$\sqrt{N}$ h-v lines	N points	N points	$\sqrt{N}$ lines
representation	2D array of N lists	N-node BST	N-node multilist	~N-node BST
cells	~N squares	N rectangles	N polygons	~N triangles
search cost	1	log N	log N	log N
extend to kD?	too many cells	easy	cells too complicated	use (k-1)D hyperplane