Approximate Voronoi Diagrams

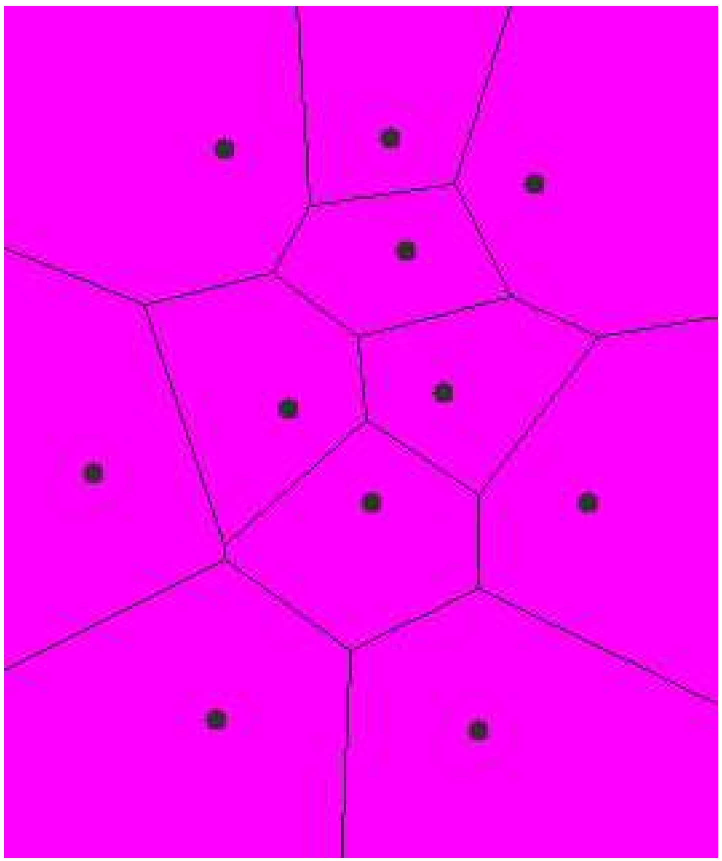
Presented by:
Guilherme Fonseca

CATS talk on 03/2005.

Based on papers by
Arya, Malamatos
and Mount

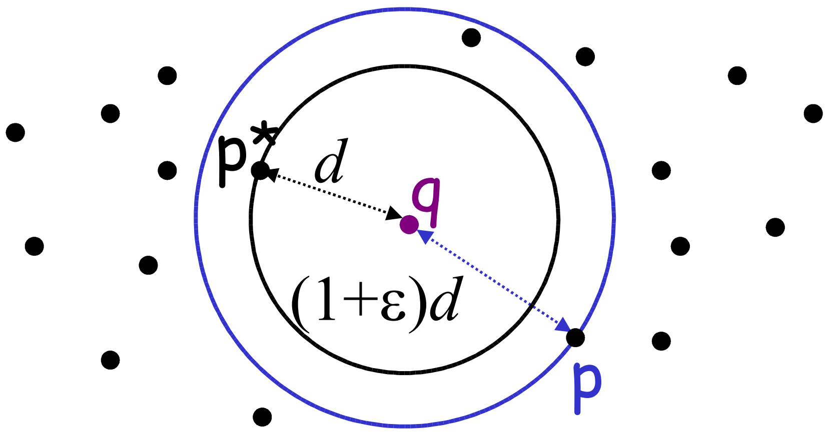
Many pictures by:
David Mount

Nearest Neighbor Queries



- Given a set *S* of points, find the closest point in *S* to a query point *q*.
- In 2 dimensions: Voronoi Diagram is very efficient!
- In $\geq$ 3 dimensions: No efficient solutions.

Approximate Nearest Neighbor

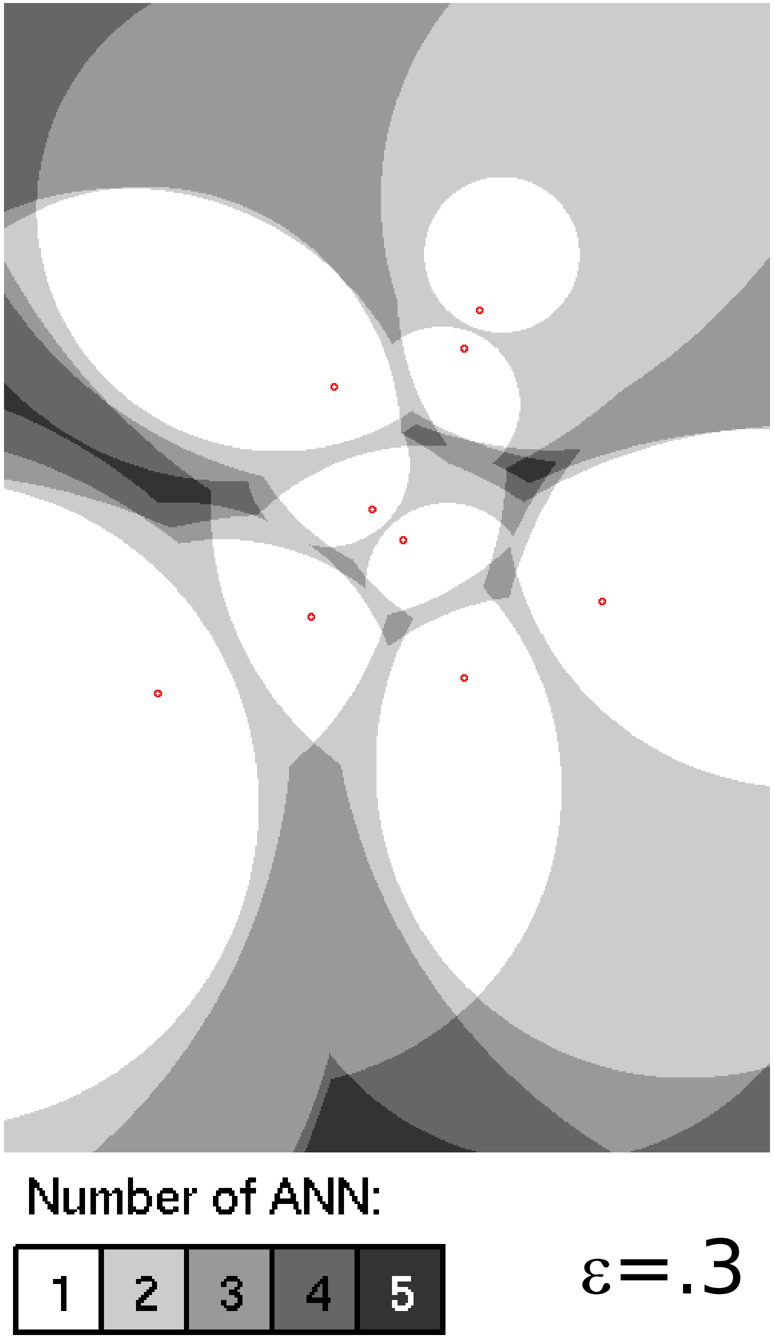


- What if we also accept a point that is not much further than the closest point?

- Given a query point q and a constant $\varepsilon < 0$, we say that p is an approximate nearest neighbor of q if:

$$\text{dist}(p,q) \leq (1+\varepsilon)\ \text{dist}(p^*,q)$$

where p^* is the actual nearest neighbor of q.

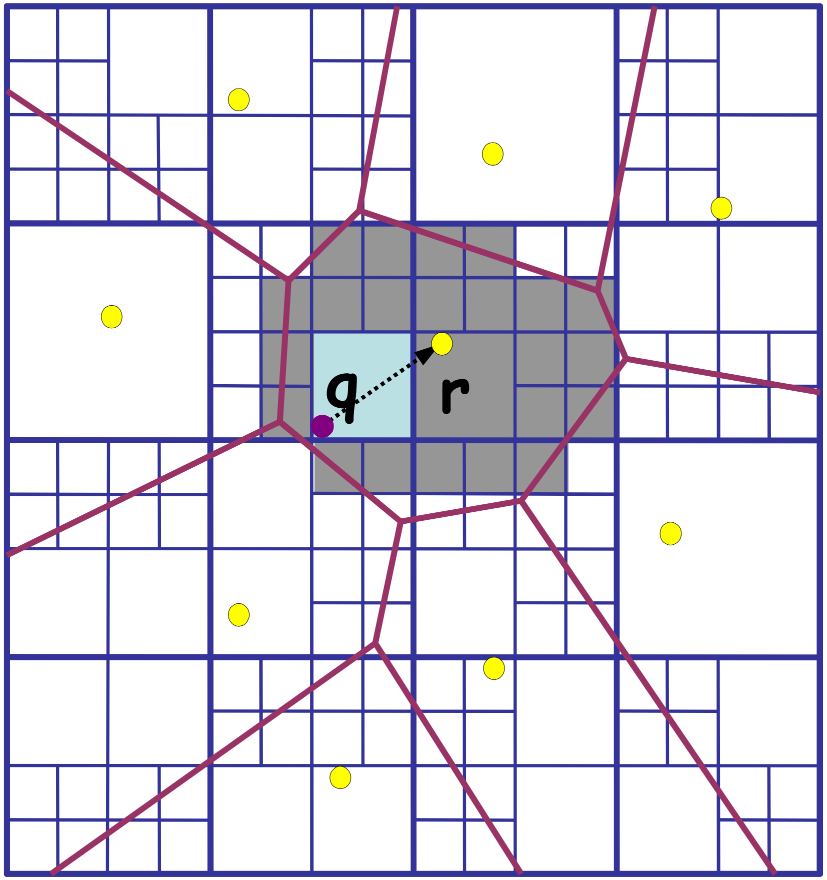


Goals

- Size: Linear in *n*
- Query: Logarithmic in *n*
- Build: Near-linear in *n*

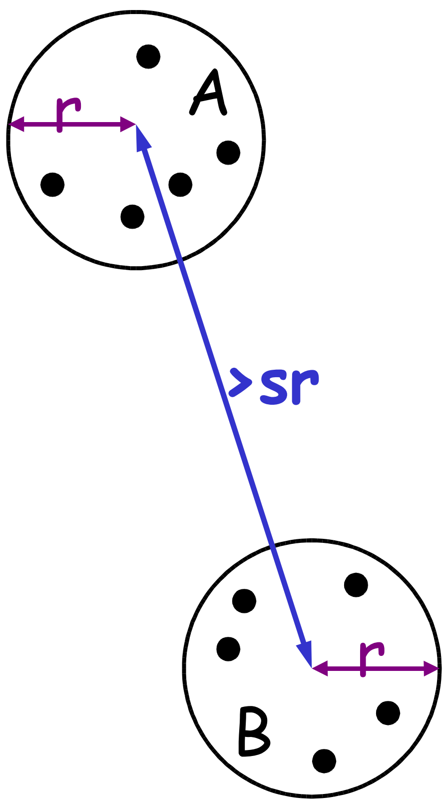
- Parameter sizes:
  - d:constant (~10)
  - 1/ε: not constant, but small (~100)
  - n: large (~10000)

Approach



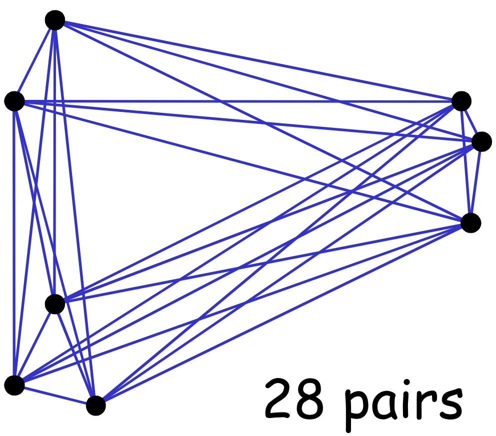
- Approximate the Voronoi Diagram with a quadtree.
- Each cell stores the approximate nearest neighbor for the corresponding region.

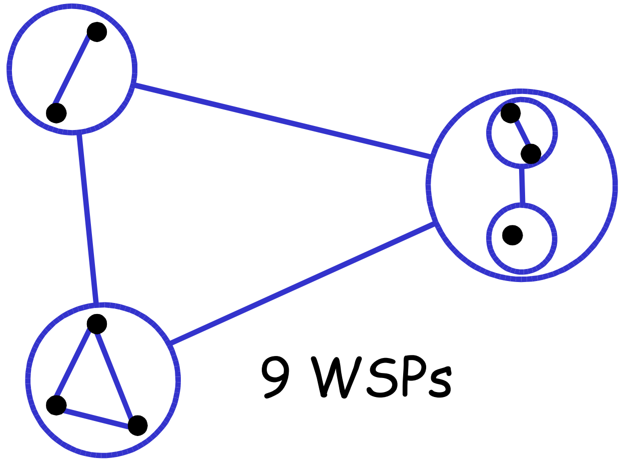
Main Tool: Well Separated Pair Decomposition



- Given a separation constant *s* (let's say *s*=8), the sets *A* and *B* are well separated if:
- Each of *A* and *B* can be enclosed by spheres of radios *r* such that the center of the spheres dist more than *sr*.

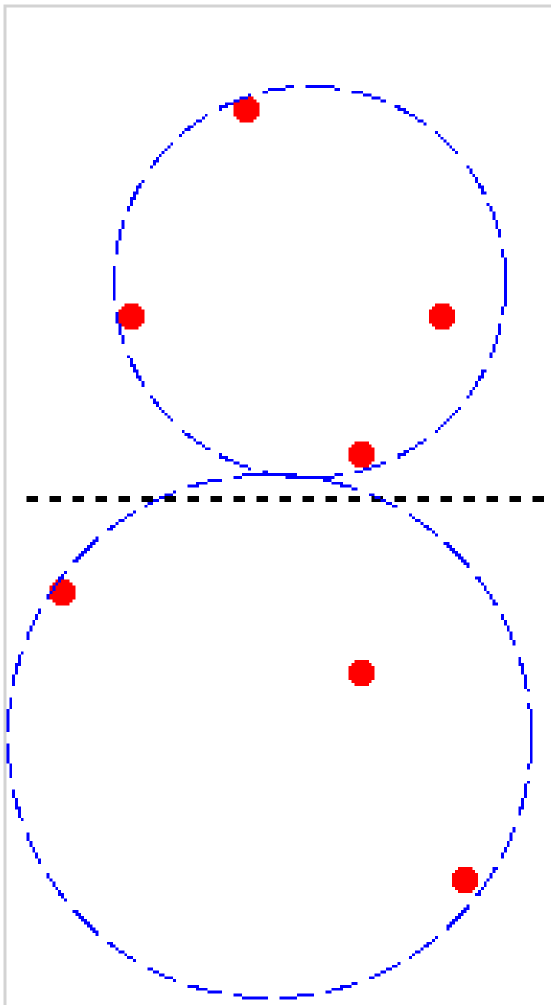
Main Tool: Well Separated Pair Decomposition





- WSPD: a set of well separated pairs such that each pair of points is separated by exactly one pair.
- Theorem: For any separation constant *s*, there is a WSPD of size O(*n*) which can be found in O(*n* log *n*) time.

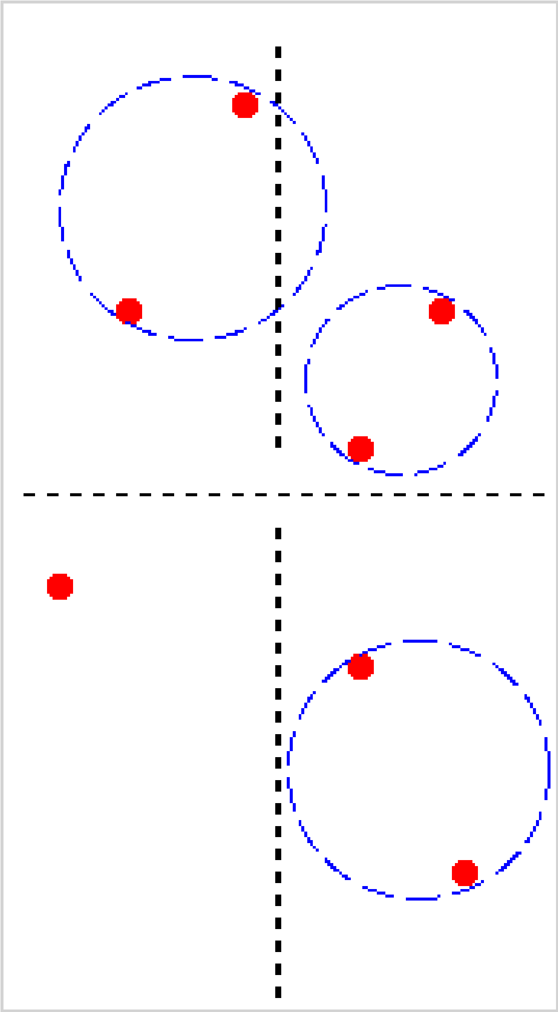
Main Tool: Well Separated Pair Decomposition



s=4

- No WSP found yet

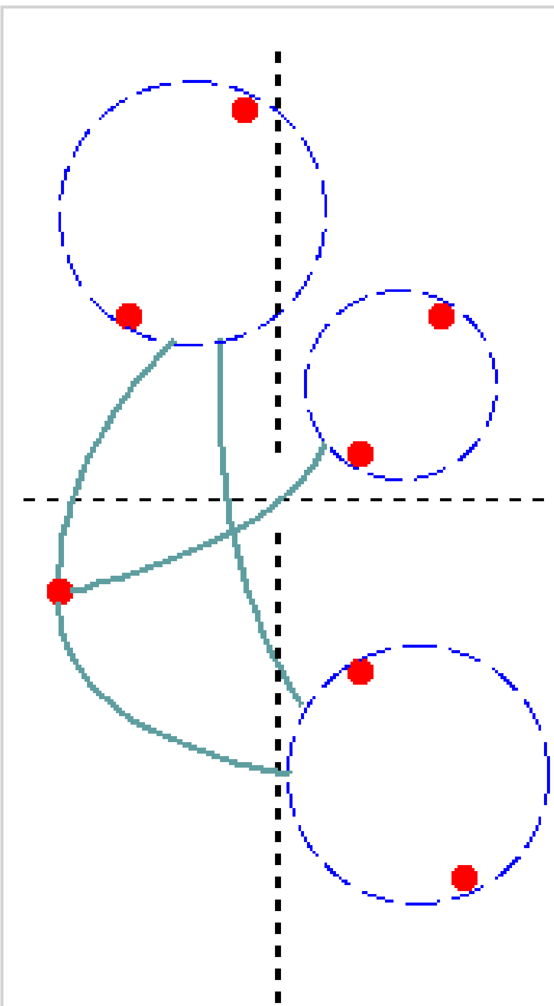
Main Tool: Well Separated Pair Decomposition



s=4

- Which sets are well separated?

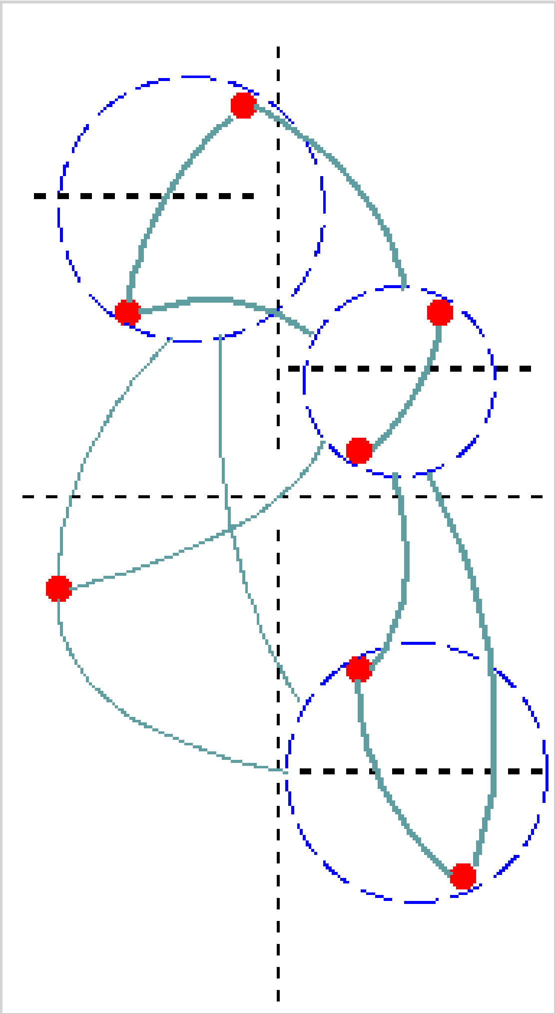
Main Tool: Well Separated Pair Decomposition



s=4

- Which sets are well separated?

Main Tool: Well Separated Pair Decomposition

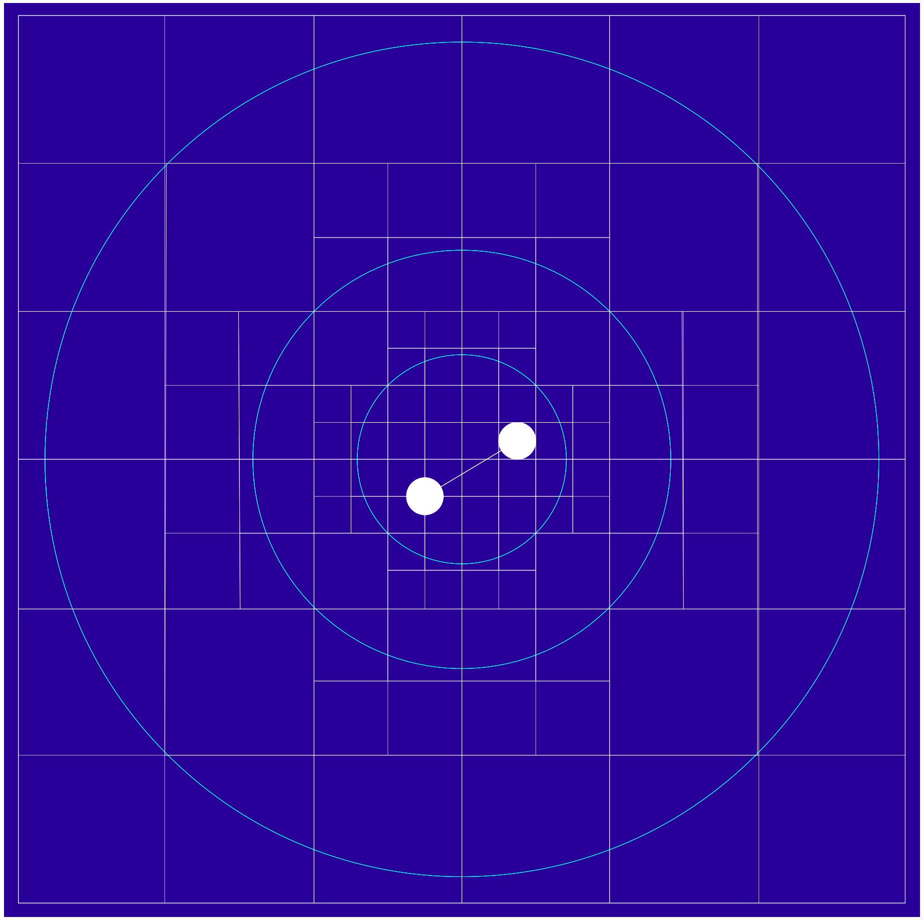


s=4

- Total: 11 WSPs

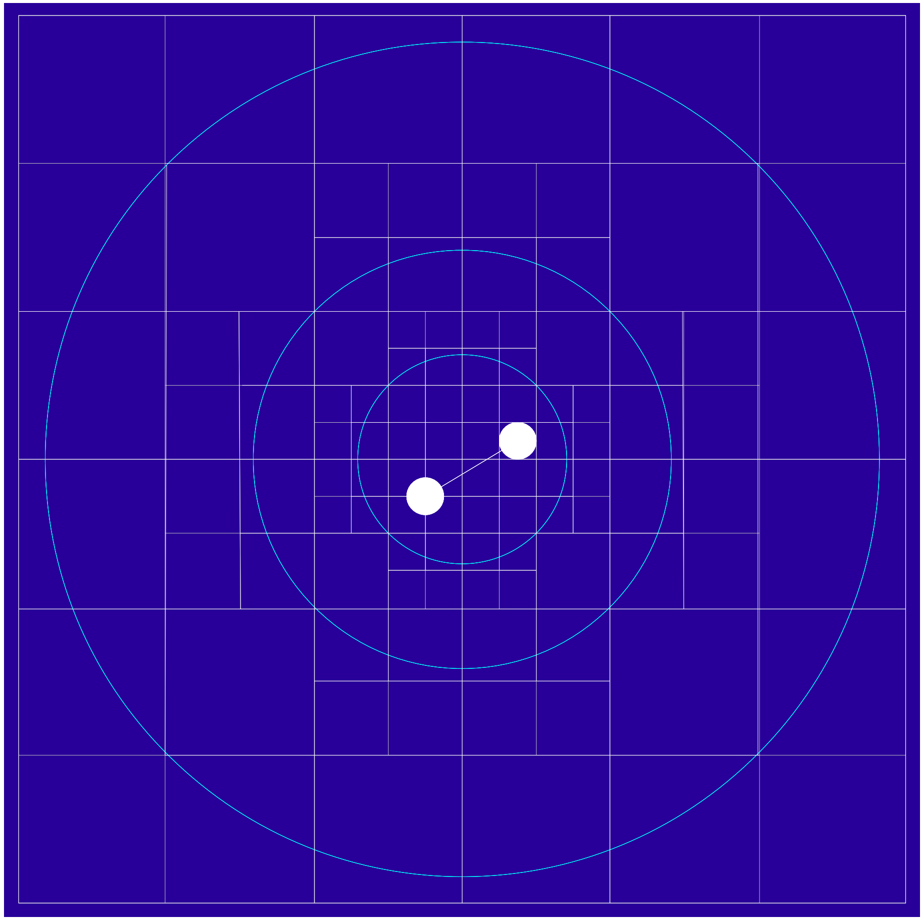
(Out of 21 pairs total.)

Construction



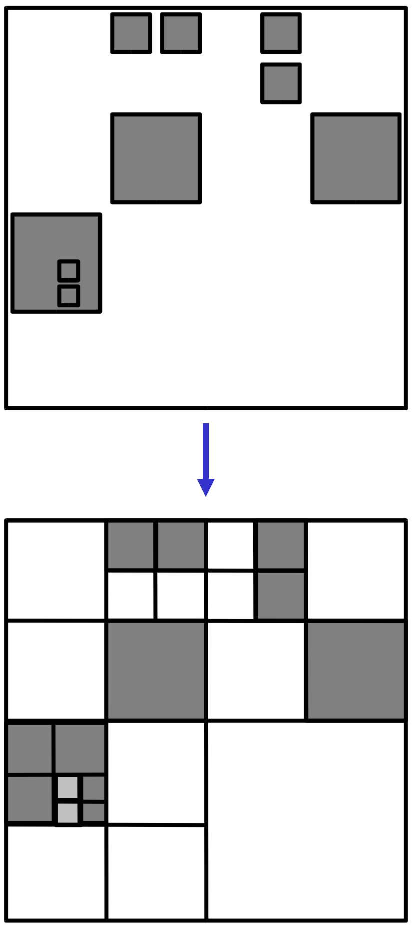
- Find a WSPD with s=8
- For each pair:
  - Create concentric balls b_i of radius r_i
  - Generate quad-tree boxes of size $f(r_i)$ intersecting ball b_i
- For each box, find a representative

Compactness

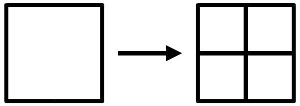
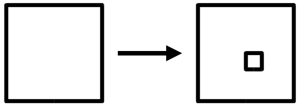


- The number of quad-tree boxes is linear in n:
  $O(n/\varepsilon^d \log (1/\varepsilon))$
- It's possible to reduce it to $O(n/\varepsilon^d)$ by making $f(r)$ quadratic instead of linear.

Another tool: BBD tree



- Given *n* quadtree boxes, we can build a tree of
  - size O(*n*)
  - height O(log *n*)

- BBD Tree Subdivision:
  - splitting: 
  - shrinking: 

Correctness

- Three cases:

  (1) $|qy'| \geq 2l/\varepsilon$

  (2) $2l/\varepsilon \geq |qy'| > l/4$

  (3) $l/4 \geq |qy'|$
- Using triangle inequality many times we show that (3) cannot happen and x is ANN(q) in (1) and (2).

q: query point

x: reported ANN(q)

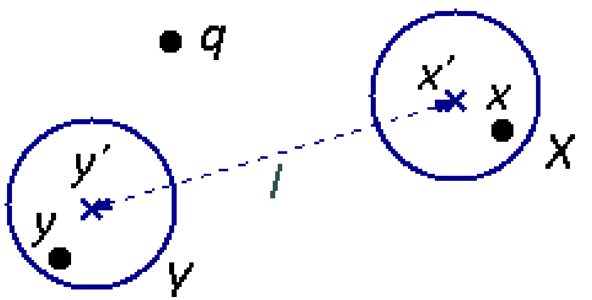
y: true NN(q)

(X,Y): WSP of x,y

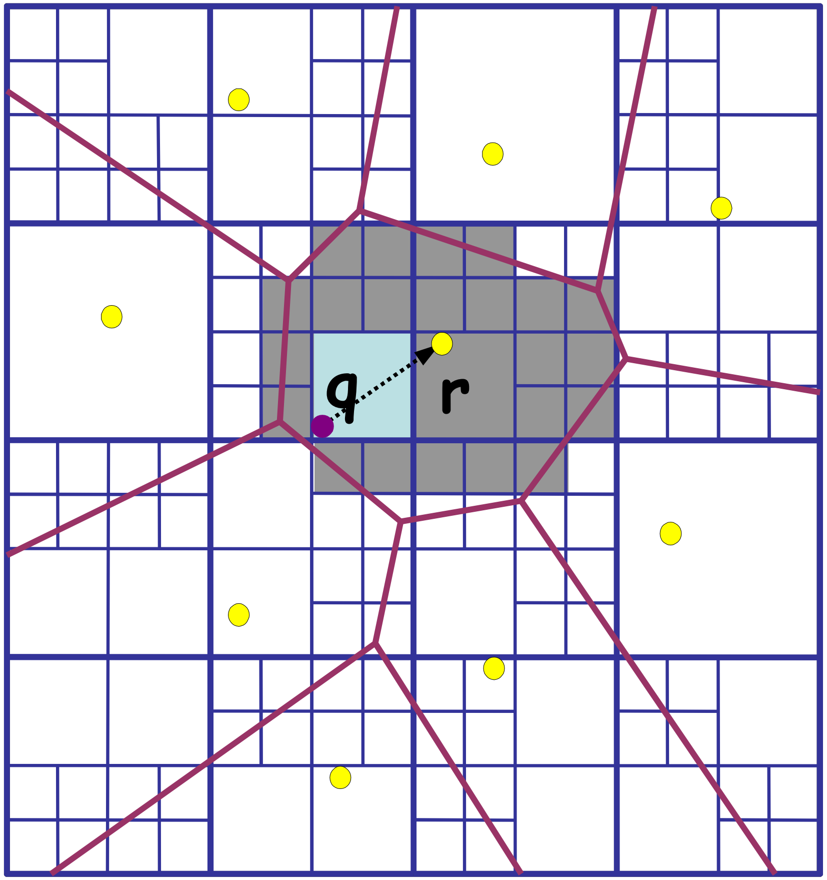
x': Center of X ball

y': Center of Y ball

l: dist(x',y') = $|xy'|$

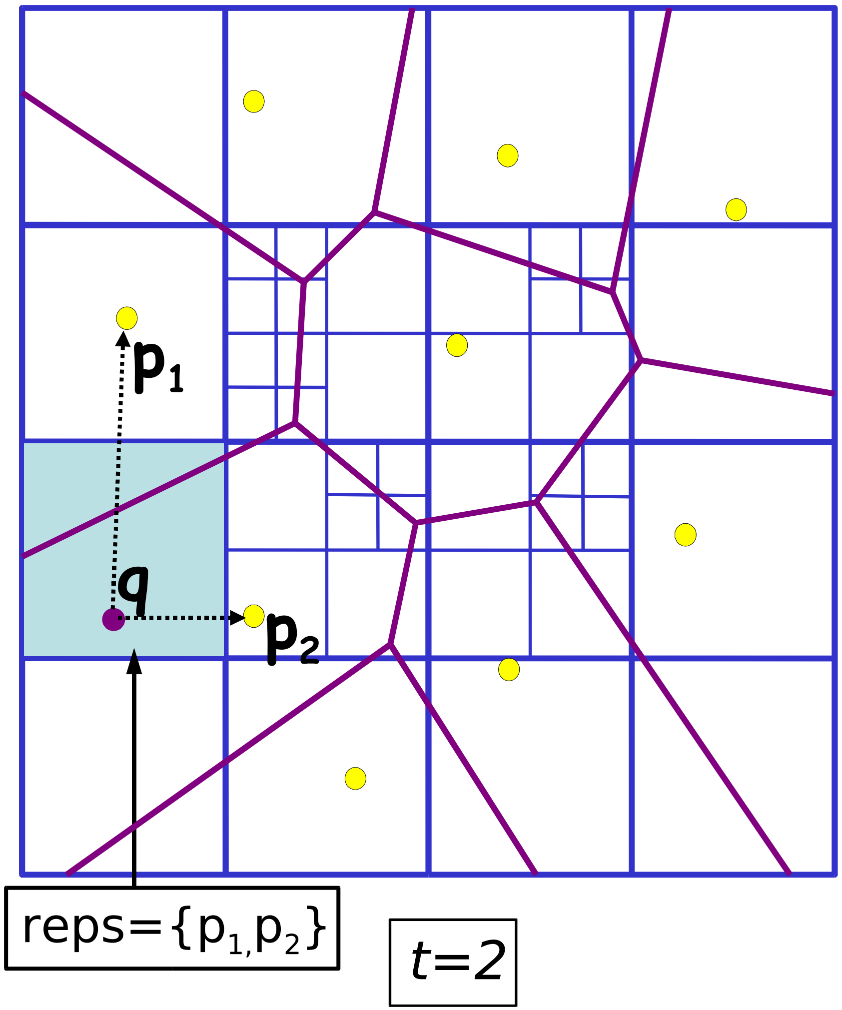


Results



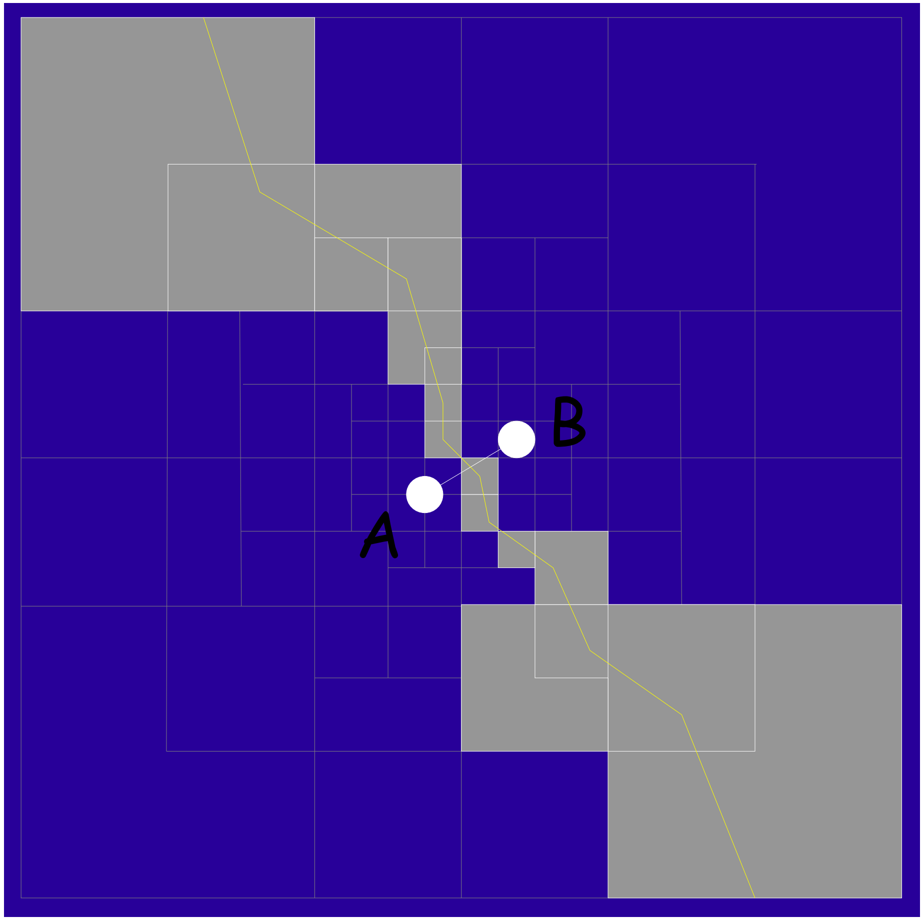
- Size:
  $O(n/\varepsilon^{d})$
- Query time:
  $O(\log n/\varepsilon)$
- Lower bound on the number of quadtree boxes:
  $\Omega(n/\varepsilon^{d-1})$

Improvement: Multiple Representatives



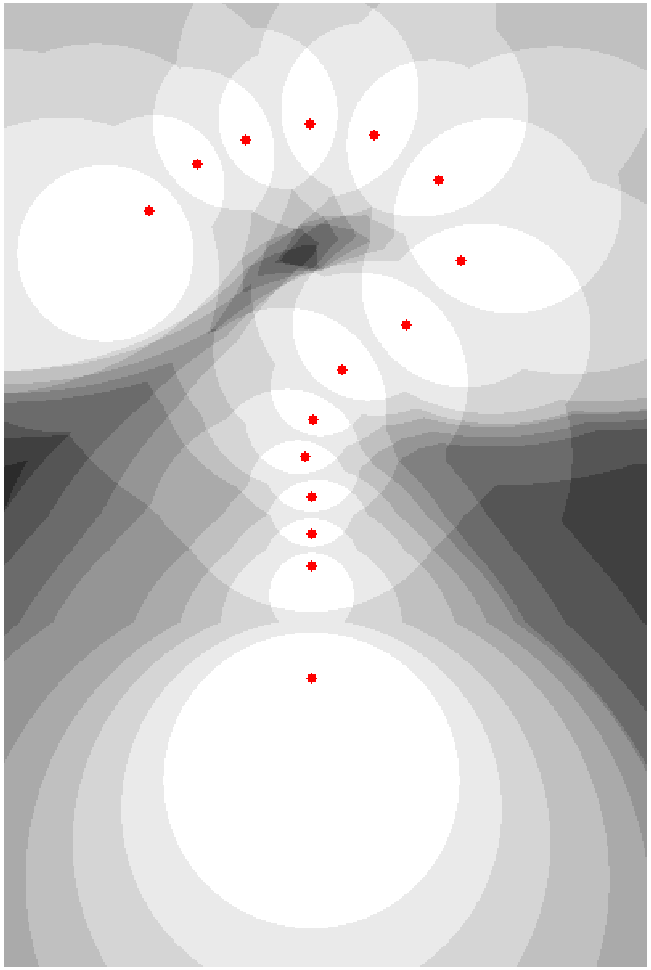
- Each cell can store t representatives. We search all of them to find the closest representative.
- If $t=O(1/\varepsilon^{(d-1)/2})$, we have:
  - Size: $O(n/\varepsilon^{(d-1)/2})$ (single rep: $O(n/\varepsilon^{d})$)
  - Query time: $O(\log n + 1/\varepsilon^{(d-1)/2})$

Improvement: Bisector Sensitivity



- When processing WSP (A,B), we only create the boxes that intersect the A-B bisector.

Open Problems



- AVDs for other problems. ex: farthest point
- AVDs for other objects?
- Dependence on ε: Must construction depend on ε?
- Further reduce size to make AVDs practical.

Thank you!