Modify Range Search to find NNs

Def Given a set $P \subset \mathbb{R}^{\cancel{2}\, d}$ and a point $q \in \mathbb{R}^{\cancel{2}\, d}$ the <u>nearest neighbor</u> of q (in P) is the point $p \in P$ s.t. $\|p-q\| \le \|p'-q\|$ for all $p' \in P$.

A first attempt:

```
NN₁(T, q)
  If T is a single node, return T.point
  Else
  { let a = NN(T.left, q)
    let b = NN(T.right, q)
  } If ||a-q|| ≤ ||b-q|| return a else return b.
```

This is Brute Force

A second Attempt

```
NN₂(T, q)
  If T is a single node, return T.point
  Else If q ∈ T.left return NN(T.left, q)
       Else return NN(T.right, q)
```

gives wrong answer (quickly!)

$NN(T: \text{tree},\ q: \text{query},\ p: \text{guess})$

If T is a leaf

$\{$ ~~return T.point~~

return $(\|q-p\| < \|q - T.\text{point}\|)\ ?\ p : T.\text{point}$ $\}$

If $(T.\text{left} \cap \text{ball}(q, \|q-p\|) \neq \emptyset)$

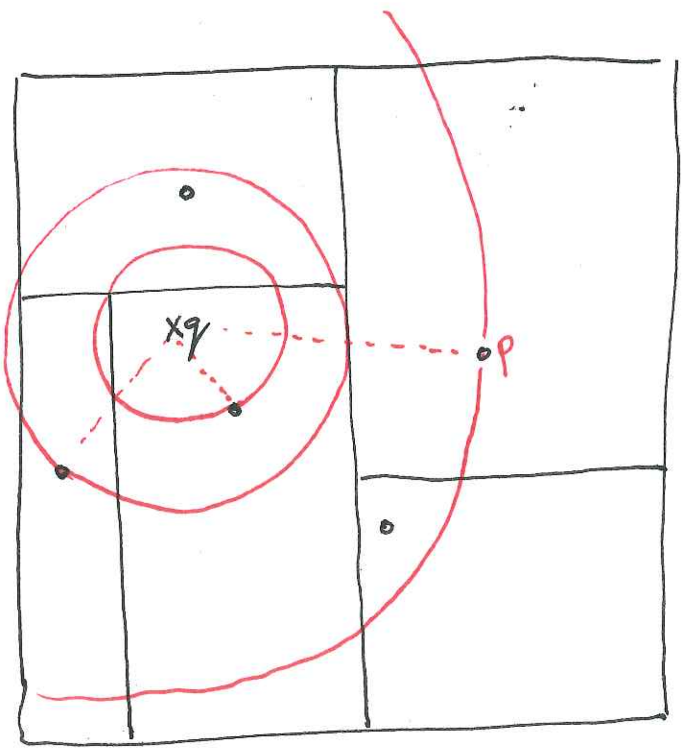
$\{ \ell = NN(T.\text{left}, q, p) \}$

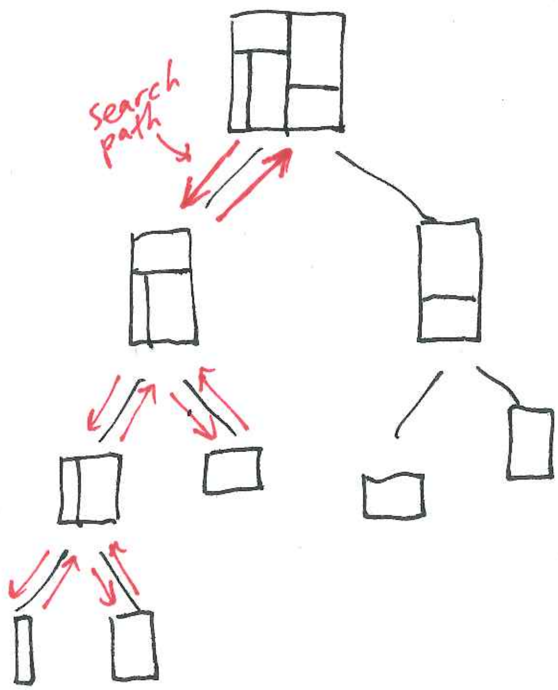
If $T.\text{right} \cap \text{ball}(q, \|q-\cancel{p}\ \ell\|) \neq \emptyset)$

$\{$ ~~r =~~ return $NN(T.\text{right}, q, \cancel{p}\ \ell) \}$

Optimization (in red) use ℓ as the new guess because it could be better than p.

~~return $(\|q-\ell\| < \|q-r\|)\ ?\ \ell : r$~~





Def Let $P \subset \mathbb{R}^2$ and let $q \in \mathbb{R}^2$.
For a constant $c \in \mathbb{R}$, a <u>c-approximate nearest neighbor</u> (or <u>c-ANN</u>) of q is a point $p \in P$ s.t. $\|p-q\| \le c\|q-p'\|$ for all $p' \in P$.

Note: It's easy to adapt our current NN code.
Replace $\text{ball}(q, \|p-q\|)$ with $\text{ball}(q, (\frac{1}{c})\|p-q\|)$

ANN(T : tree, q : query)

If T is a leaf, return T.point

For $i = 1$ to 4

{If $q \in$ T.child(i)

{p = ANN(T.child(i), q)}

}

If p == nil, return whatever(T) any point will do.

else return p.

Assume $P \subset \underbrace{[0,1] \times [0,1]}_{\text{unit square}}$

Level i of the QT has squares with sidelength $\frac{1}{2^i}$ (so root is level 0).

The <u>scale</u> at level i is $\frac{1}{2^i}$

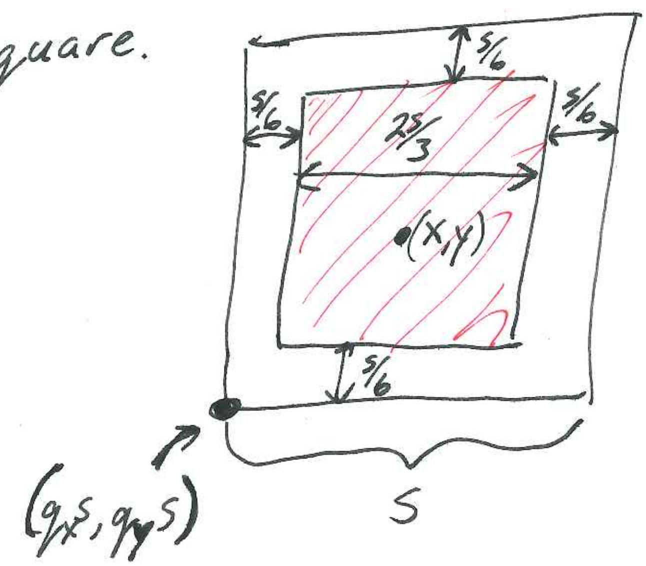
A square containing a point x at scale s is denoted square(s, x)

<u>Def</u> A point (x,y) is <u>central</u> in Square$(s, (x,y))$ if there are integers ~~q~~ q_x, q_y s.t.

$$\frac{s}{6} \leq x - q_x s \cancel{\leq} \frac{5s}{6}$$

and
$$\frac{s}{6} \leq y - q_y s \cancel{\leq} \frac{5s}{6}$$

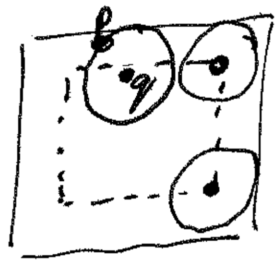
In other words (x,y) is in the middle $\frac{2}{3}$ of the square.



Let q be the query and let $p = NN(q)$.

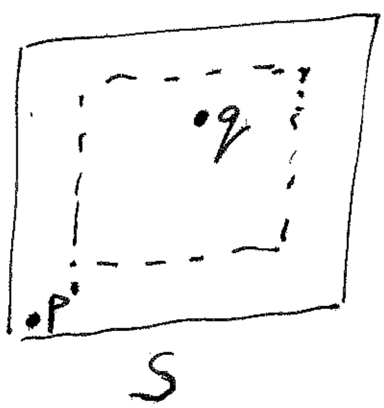
Lem 1 If q is central in square(s,q) for $s > 6\|p-q\|$ then $p \in$ square(s,q).

pf sketch



Lem 2 If q is central in square(s,q) and square(s,q) contains at least one point of P then $\|p-q\| \le s\sqrt{2}$ and $\|q - ANN(QT,q)\| \le s\sqrt{2}$

pf sketch



Let $p' = ANN(QT,q) \in$ square(s,q).
$\|p-q\| \le \|p'-q\| \le s\sqrt{2}$.

☆ It would be nice if q is central in $\text{Square}(s, q)$ where

$$6\|pq\| \leq s < 12\|pq\|$$

$\boxed{\text{Why?}}$

By Lem 1, $p \in \text{square}(s, q)$

So, by Lem 2, $\|q - ANN(QT, q)\| \leq s\sqrt{2} < 12\sqrt{2}\|pq\|$.

Thus, by def^n, $ANN(QT, q)$ is a $12\sqrt{2}$-ANN.

not great, but it is constant.

Idea: Shift the Coordinate System

Store 3 QTs. Search in parallel.
(search 3 times and pick best answer)

Let $v_i = (\frac{i}{3}, \frac{i}{3})$ for $i \in \{0, 1, 2\}$

$P_i = \{p + v_i : p \in P\}$ (so, $P_0 = P$)

QT_i is the quadtree for P_i

New Algorithm:

For $i = 0 \ldots 2$
  answer[i] = ANN(QT_i, $q + v_i$)

return best answer.

Claim: For any $p, q \in [0,1]^2$, there is a shift $i \in \{0,1,2\}$ s.t. q is central in square(s, q) where

$$6\|pq\| \le s < 12\|pq\|$$

this should blow your mind a little.