3/2/21

Linear time algorithm for closest pair

$P = \{p_1, \dots, p_n\}$

$p_i = (x_i, y_i)$

$\underline{cp(P)} = \min_{i<j} \|p_i - p_j\|$

1dim

Rabin 1976 ⟵

- floor function

  $\lfloor x \rfloor$ = integer part of x.

- Randomization
- Hashing.

Problem

π: random permutation of $1, \dots, n$

$\pi(1), \pi(2)$

$m_i = \min(\pi(1), \pi(2), \dots, \pi(i))$

4 2 3, 1

$m_2 = 2$

$m_1, m_2, m_3, \dots \quad m_n$

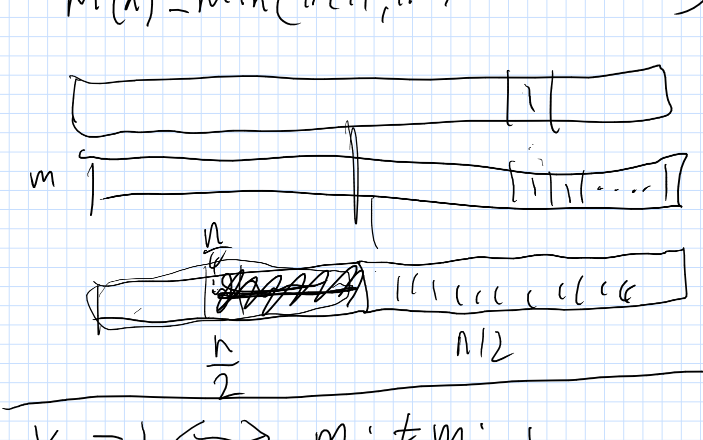
$\pi(1), \pi \quad \dots \quad \pi(n)$

Q: How many different numbers are there in $m_1, m_2, m_3, \dots$

M = random variable

Q: $E[M] = ?$

$m(n) = \min(\pi(1), \pi(2) \dots \pi(n)\} = 1$



$X_i = 1 \iff m_i \neq m_{i-1}$

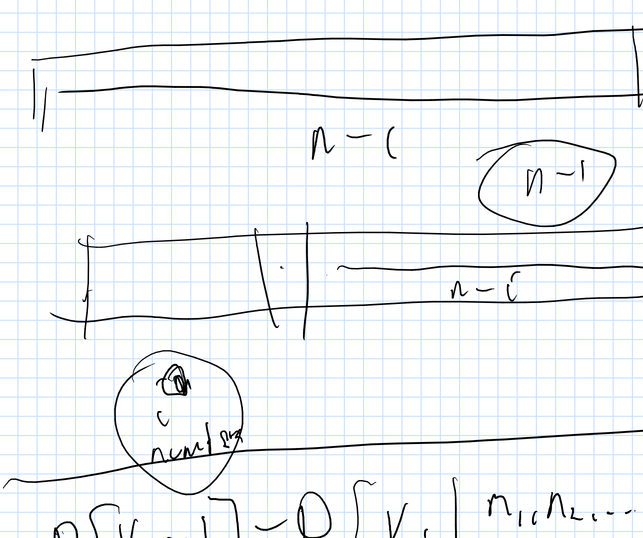
$\pi(i) < \pi(1), \pi(2), \dots \pi(i-1)$

$P[X_i = 1]$

$P[X_n = 1] = \frac{1}{n}$

$P[X_i = 1] = \frac{1}{i}$

Backward analysis



$$P[X_i = 1] = P\left[X_i \,\middle|\, \begin{array}{l} n_1, n_2, \dots n_i \\ \text{are the first} \\ i \text{ numbers} \end{array}\right]$$

$$= P\left[X_i = 1 \,\middle|\, \begin{array}{l} n_1 = 1 \\ n_2 = 2 \\ \vdots \\ n_i = i \end{array}\right] = \frac{1}{i}$$

$$E[M] = E\left[\sum X_i\right] = \sum E[X_i]$$

$$= \sum_{i=1}^{n} \frac{1}{i} = O(\log n)$$

- verification

  $x_{i+1} > m_i \qquad n - \log n$

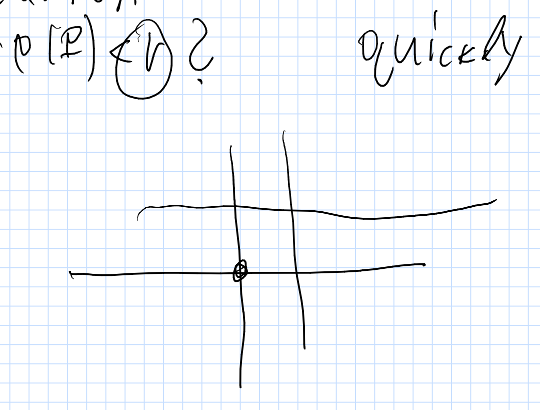
- update

  $m_i \leftarrow x_i \qquad O(\log n)$

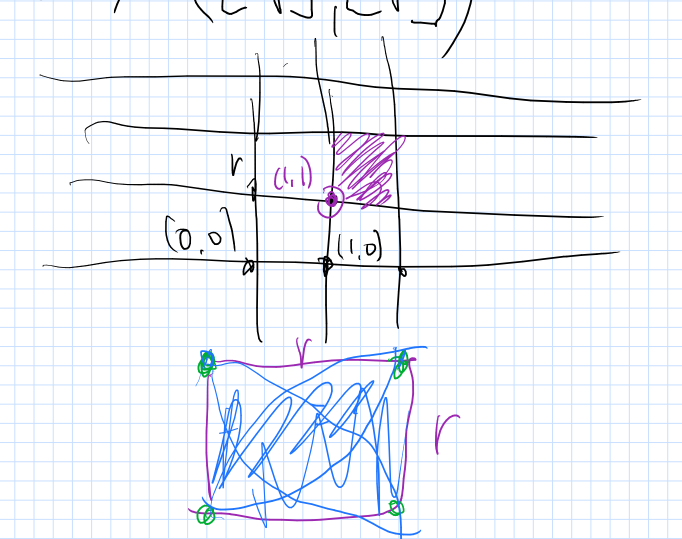
$P = \{p_1, p_2 \dots p_n\}$

Verification

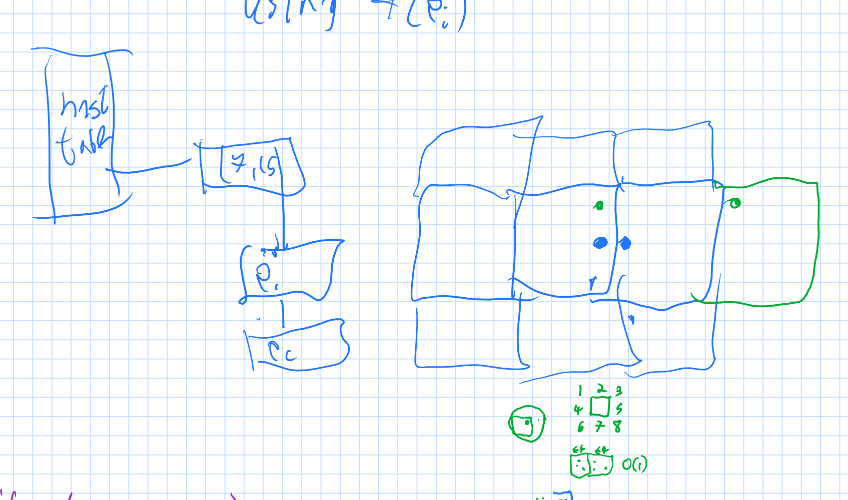
$cp(P) < r$? quickly!



$$f(x, y) = \left(\left\lfloor \frac{x}{r} \right\rfloor, \left\lfloor \frac{y}{r} \right\rfloor\right)$$



$p_i \quad f(p_i)$ store p_i in hash table using $f(p_i)$



Verification $cp(P) \leq r$ can be done in $O(n)$ time.

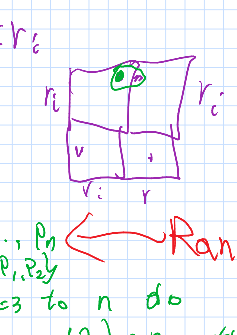
$r_i = cp(p_1, p_2, \dots, p_i)$

$\underline{P_i} = \{p_1, p_2, \dots, p_i\}$

p_{i+1} — new point

$r_{i+1} = cp(\underline{P}_{i+1} = P_i \cup \{p_{i+1}\}) < r_i$

$r_{i+1} < r_i$



$p_1, p_2, \dots, p_n$ ⟵ Random permutation of the points

$r_2 = cp(\{p_1, p_2\})$

For $i = 3$ to n do

  if $cp(P_i) < r_{i-1}$ then

    $r_i \leftarrow$ get from verification test

    rebuild the grid with side length r_i] $O(i)$ ⟸

  else

    store p_i in the grid(r_{i-1})

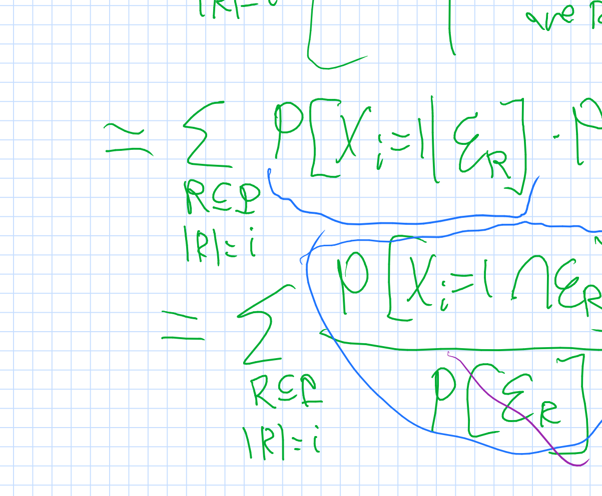
    $r_i \leftarrow r_{i-1}$] $O(1)$

$\sum_{i=1}^{n} O(i) = O(n^2)$ ⟵

$E[\text{running time}] = ?$

$X_i = 1 \iff r_i < r_{i-1}$

$$P\left[X_i \,\middle|\, \begin{array}{l} q_1, q_2, \dots, q_i \\ \text{are the first} \\ i \text{ points of the permutation} \end{array}\right]$$



$$P[X_i = 1] = \sum_{\substack{R \subseteq P \\ |R| = i}} P\left[X_i = 1 \,\middle|\, \begin{array}{l} \text{the first} \\ i \text{ points} \\ \text{in the} \\ \text{permutation} \\ \text{are } R \end{array}\right] \quad (\mathcal{E}_R)$$

$$= \sum_{\substack{R \subseteq P \\ |R| = i}} P[X_i = 1 \mid \mathcal{E}_R] \cdot P[\mathcal{E}_R]$$

$$= \sum_{\substack{R \subseteq P \\ |R| = i}} \frac{P[X_i = 1 \cap \mathcal{E}_R]}{P[\mathcal{E}_R]} P[\mathcal{E}_R]$$

$$= \sum_{R \subseteq P} P[X_i = 1 \cap \mathcal{E}_R] = P[X_i = 1]$$

$$P[X_i = 1] = \sum_R P[X_i = 1 \mid \mathcal{E}_R] P[\mathcal{E}_R]$$

$$\leq \sum_{\substack{R \subseteq P \\ |R| = i}} \frac{2}{i} P[\mathcal{E}_R] = \frac{2}{i} \sum_{\substack{R \subseteq P \\ |R| = i}} P[\mathcal{E}_R] = \frac{2}{i}.$$

$P[X_i = 1] \leq \frac{2}{i}$

W_i = amount of work done by the algorithm in ith iteration

$$W_i = \begin{cases} O(1) & r_i = r_{i-1} \\ O(1) + O(i) & r_i < r_{i-1} \end{cases}$$

$W_i = 1 + X_i \cdot i$

$E[W_i] = O(E[1 + X_i \cdot i])$

$= O(1 + E[X_i] \cdot i)$

$= O(1 + P[X_i = 1] \cdot i)$

$= O(1 + \frac{2}{i} \cdot i) = O(3) = O(1)$

$$E[RT] = E\left[\sum_{i=1}^{n} W_i\right]$$

$$= \sum_{i=1}^{n} E[W_i] = \sum_{i=1}^{n} O(1) = O(n).$$

$\lfloor x \rfloor$