Quick sort

Rispetto al selection sort e all'insertion sort, il quick sort usa il paradigma **DIVIDE ET IMPERA**, ossia divide il problema in sottoproblemi che può risolvere e poi li usa per ottenere la soluzione finale (fase di "impero").

Idea: consideriamo ordinare una lista in tre porzioni:

$$\underbrace{\boxed{\le px}}_{\text{cose min. o ug. di } px} \quad \underbrace{\boxed{px}}_{\text{pivot}} \quad \underbrace{\boxed{\ge px}}_{\text{cose magg. o ug. di } px} \Big\}\ n \text{ elem.}$$

Questo possiamo farlo facilmente prendendo due indici i, j, spostando px in fondo e:

- a ogni step facciamo andare avanti i e j (i verso destra da $A[0]$, j verso sin da $\underbrace{A[n-2]}_{A[n-1] = px}$) fino a incontrare i t.c. $A[i] > px$ e j t.c. $A[j] < px$ e scambio $A[i]$ e $A[j]$.

- mi fermo non appena $i > j$ e risposto px.

Possiamo quindi a ogni passo scegliere un pivot px, partizionare l'array come appena fatto (fase di DIVIDE) e applicare ricorsiv. l'alg. sulle parti estreme, incollandole alla fine (fase di IMPERA)

Il partizionamento effettuato in qst modo è $O(n)$ — con n lunghezza dell'array.

Il caso pessimo è $O(n^2)$, se i pivot presi sono quasi sempre + massimi/minimi dei segmenti (diventa moralmente un selection sort). Il caso medio è però $O(n \cdot \log n)$.

Scelta del pivot

<u>Non</u> c'è un modo efficace e strettamente deterministico di scegliere il pivot. Possiamo quindi provare a scegliere il pivot randomicamente (Quicksort randomizz.) supponendo di avere una funzione random(a, b) che restituisce un num. in $[a, b]$ con prob. $1/(b-a+1)$. Questa scelta ci garantirà ancora un caso medio $O(n \cdot \log n)$, ma su una classe "più probabile" di problemi.

Sia X il numero di confronti nel QS randomizz. Allora X è una v.a. t.c.

$$X = \sum_{i<j} X_{ij}, \quad \text{sorted}(A) = (z_i)$$

$$X_{ij} = \begin{cases} 0 & z_i, z_j \text{ \underline{non} confront.} \\ 1 & \text{altrim.} \end{cases}$$

Infatti i e j vengono confrontati al più una volta — e uno deve essere il pivot dell'altro.

Inoltre $z_i, z_j \in$ partizione $\Rightarrow$
$\rightarrow z_i, z_{i+1}, \dots, z_{j-1}, z_j \in$ partiz.

Da questo:

$$E[X] = \sum_{i<j} E[X_{ij}] = \sum_{i<j} P(X_{ij} = 1) \overset{(*)}{\leq}$$

$$\leq \sum_{i<j} \frac{2}{j-i+1} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} =$$

$$= \sum_{i=1}^{n-1} \sum_{k=2}^{n-i+1} \frac{2}{k} \sim n \cdot \underbrace{\log(n)}$$

ciclo esterno

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

(*) è dovuto al fatto che $P(X_{ij} = 1)$ è dunque la prob. che z_i o z_j siano pivot dell'altro e contenendo la partiz. $z_i, z_{i+1}, \dots, z_{j-1}, z_j$ ed essendo P unif. ed è quindi $\leq P(z_i$ pivot di $z_j) +$ $+ P(z_j$ pivot di $z_i) \leq$

$$\leq \frac{1}{j-i+1} + \frac{1}{j-i+1}.$$