

MAJORITY ELEMENT

• DATO UNO STREAM Σ , RESTITUIRE L'ELEMENTO DI MAGGIORANZA SE ESISTE

• SE $|\Sigma| = n$, x E' L'ELEM. MAGGIORANZA DI Σ
SE $f(x) \equiv \text{FREQU. DI } x \geq \lfloor \frac{n}{2} \rfloor + 1$ $\left[f(x) = \text{FREQUENZA DI } x \text{ IN } \Sigma \right]$

• L'ELEM. MAGG. POTREBBE NON ESISTERE
ESEMPIO $\Sigma = 1, 2, 3, 4, 4, 5, 5$

ALGORITMO DI BOYER-MOORE (1981)

$m = \text{NULL}$, $i = 0$

FOR EACH x IN Σ

IF $i = 0$

$i = 1$

$m = x$

ELSE IF $m = x$

$i = i + 1$

ELSE

$i = i - 1$

RETURN m

TEOREMA: SE ESISTE UN ELEM. MAGGIORANZA, $\hat{x}$ IN S ,
ALLORA L'ALGORITMO DI BOYER-MOORE RESTITUISCE $\hat{x}$

[NOTA: SE ELEM. MAGGIORANZA NON ESISTE, L'ALGORITMO RITORNA
CONVINCENDO UN ELEMENTO DI S , i]

PROVA PER ASSUNDO, SIA OUTPUT DELL'ALGORITMO $\neq \hat{x}$
OVVERO $\exists x \in S, x \neq \hat{x}$. SIA $\hat{x}$ IL VALORE FINALE DEL
CONTATORE i ALLA FINE DI S

$$0 \leq i = n_x - n_{\neq x}$$

DOVE: $\rightarrow n_x = \#$ DI OCCORRENZE DI x DA QUANDO i VIENE
IMPOSTATO AD x PER L'ULTIMA VOLTA.

$\rightarrow n_{\neq x} = \#$ DI OCCORRENZE $\neq x$ DA QUANDO i VIENE
IMPOSTATO AD x PER L'ULTIMA VOLTA

$$0 \leq i = n_x - n_{\neq x} \leq f(x) - \sum_{l \neq x} f(l) = f(x) - f(\hat{x}) - \sum_{\substack{l \neq x \\ l \neq \hat{x}}} f(l) \\ \leq f(x) - f(\hat{x}) + \sum_{\substack{l \neq x \\ l \neq \hat{x}}} f(l) =$$

$$\leq \sum_{l \neq \hat{x}} f(l) - f(\hat{x}) \leq n - f(\hat{x}) - f(\hat{x}) = n - 2f(\hat{x}) =$$

$$= n - 2 \left(\left\lceil \frac{n}{2} \right\rceil - 1 \right) = \begin{cases} n \text{ pari} & 0 \leq i \leq -2 \rightarrow \text{CONTRADDIZIONE} \\ n \text{ dispari} & 0 \leq i \leq n - \frac{(n-2)}{2} - 2 = -1 \rightarrow \text{CONTRADDIZIONE} \end{cases}$$

ABBIAMO UNA CONTRADDIZIONE PERCHÉ ABBIAMO
IPOTIZZATO $m \neq x$

Quindi $m = x$ e l'alg. ritorna l'elem. magg.



Al termine della prima passata, non possiamo
capire se m è elem. maggioranza o no.

Effettuiamo una seconda passata.

Nella seconda passata contiamo la freq. di
 $m = x$

↳ se $f(m) \geq \lfloor \frac{n}{2} \rfloor + 1 \rightarrow m$ è elem. maggioranza

↳ se $f(m) < \lfloor \frac{n}{2} \rfloor + 1 \rightarrow$ non esiste maggioranza
e ritorna null.

Quindi esiste un algoritmo di streaming
che risolve esattamente il problema dell'elem.
di maggioranza

↳ $P = 2$ (2 passate)

↳ $S = O(1)$ word (spazio)

↳ $T = O(1)$ tempo per elemento