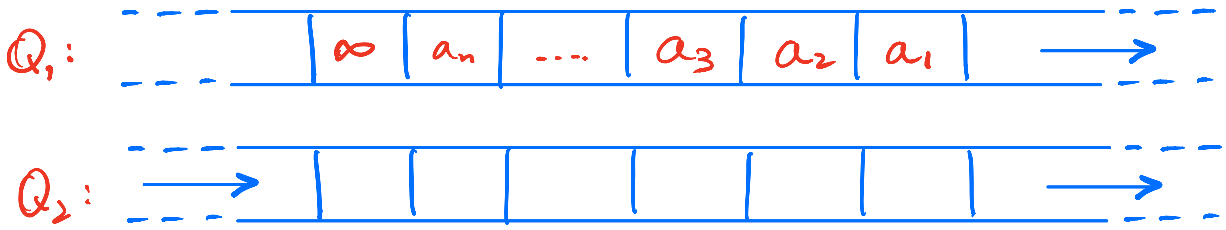
ALGORITMO DI HUFFMAN CON PREPROCESSING

- NEL CASO IN CUI I CARATTERI FOSSERO GIA' ORDINATI IN BASE ALLE FREQUENZE, NON E' NECESSARIA UNA CODA DI PRIORITA', MA BASTA UNA SEMPLICE CODA.
- CIO' CONSENTE ESTRAZIONI IN TEMPO COSTANTE PIUTTOSTO CHE LOGARITMICO E, CONSEGUENTEMENTE LA COMPLESSITA' DELLA RISULTANTE IMPLEMENTAZIONE DELL'ALGORITMO DI HUFFMAN RISULTERA' ESSERE LINEARE, A MENO DEL TEMPO DI PREPROCESSING NECESSARIO PER ORDINARE I CARATTERI IN BASE ALLE FREQUENZE
- SI UTILIZZANO DUE CODE Q_1 E Q_2 (MA SAREBBE ANCHE SUFFICIENTE UTILIZZARE UNO STACK ED UNA CODA)

INIZIALIZZAZIONE

- INIZIALMENTE LA CODA Q_1 CONTIENE TUTTI I CARATTERI/FOGLIE IN ORDINE NON DECRESCENTE DI FREQUENZA, PIÙ UN NODO SENTINELLA ∞ CON FREQUENZA "INFINITA" ($a_{n+1} = \infty$). INVECE LA CODA Q_2 E' INIZIALMENTE VUOTA.



(CON $0 \le f(a_1) \le f(a_2) \le \dots \le f(a_{n-1}) \le f(a_n)$, $n \ge 2$)

ALGORITMO

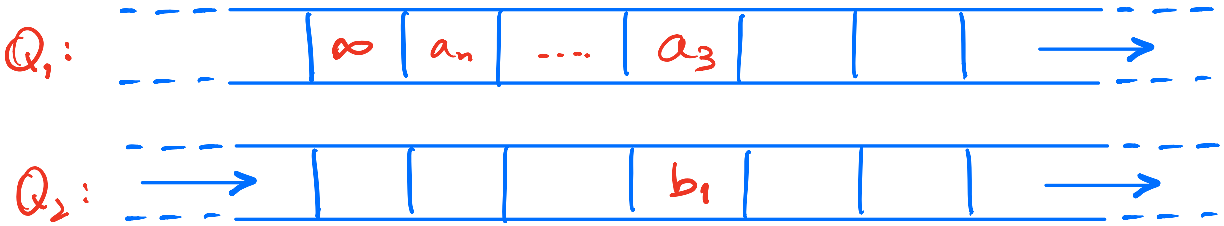
- SI ALLOCHI UN NODO INTERNO z

$$left(z) := x' := DEQUEUE(Q_1)$$

$$right(z) := x'' := DEQUEUE(Q_1)$$

$$f(z) := f(x') + f(x'')$$

$$ENQUEUE(Q_2, z)$$



(CON $f(b_1) = f(a_1) + f(a_2)$)

- QUINDI PER $(n-2)$ VOLTE SI RIPETA IL SEGUENTE PASSO:

IF $TOP(Q_1) < TOP(Q_2)$ **THEN**

  $x' := DEQUEUE(Q_1)$

  IF $TOP(Q_1) < TOP(Q_2)$ **THEN**

    $x'' := DEQUEUE(Q_1)$

  ELSE

    $x'' := DEQUEUE(Q_2)$

ELSE

  $x' := DEQUEUE(Q_2)$

  IF $IsEmpty(Q_2)$ **OR** $TOP(Q_1) < TOP(Q_2)$ **THEN**

    $x'' := DEQUEUE(Q_1)$

  ELSE

    $x'' := DEQUEUE(Q_2)$

DETERMINA I DUE MINIMI $x' \leq x''$ IN $Q_1 \cup Q_2$

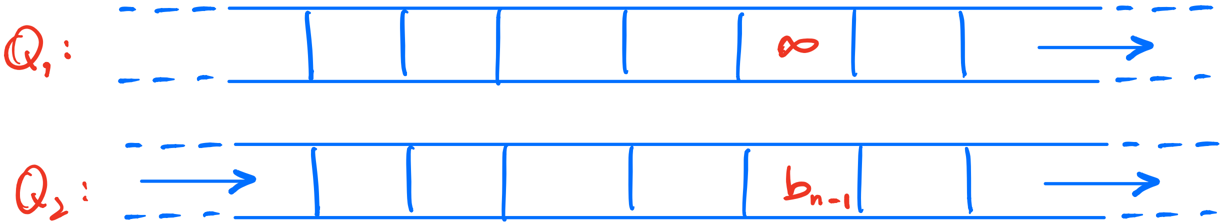
- SI ALLOCHI UN NODO INTERNO z

$left(z) := x'$; $right(z) := x''$

$f(z) := f(x') + f(x'')$

$ENQUEUE(Q_2, z)$

CONFIGURAZIONE FINALE



b_{n-1} E' LA RADICE DELL'ALBERO DI DECODIFICA

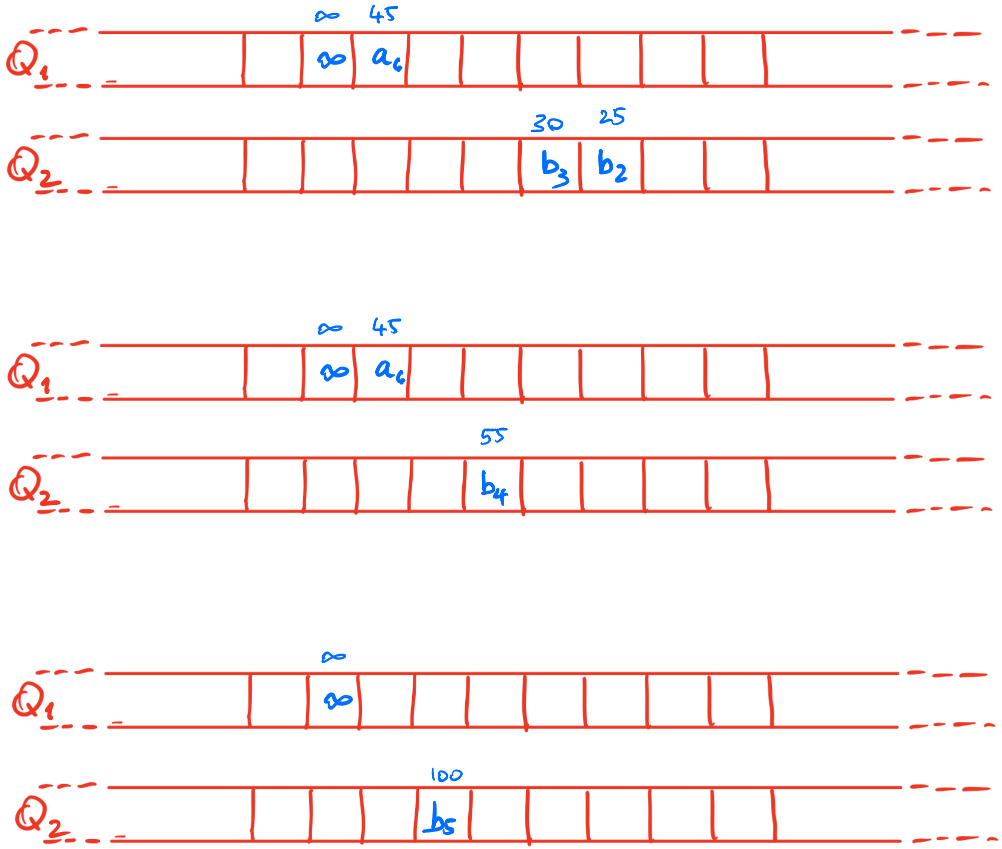
ESEMPIO

$$A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$$

$f(a_1) = 5$ $\quad$ $f(a_3) = 12$ $\quad$ $f(a_5) = 16$

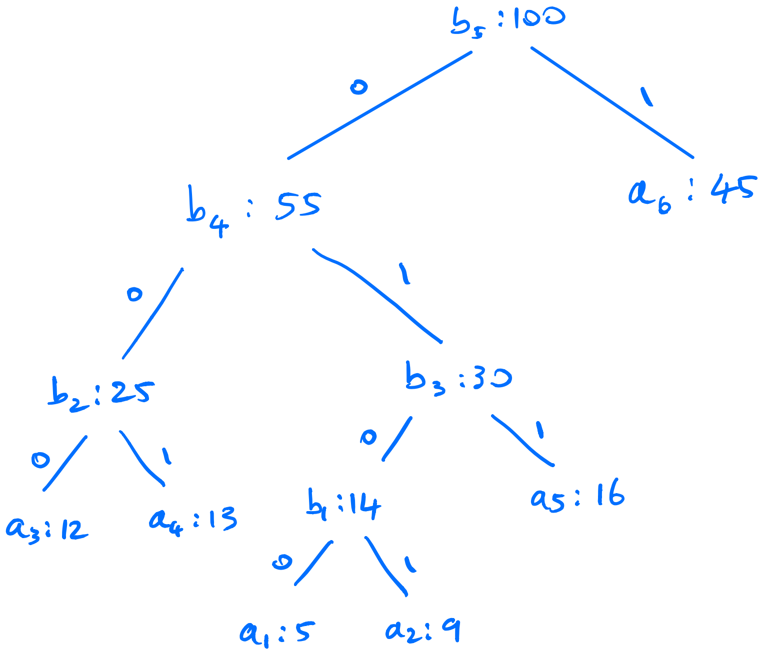
$f(a_2) = 9$ $\quad$ $f(a_4) = 13$ $\quad$ $f(a_6) = 45$

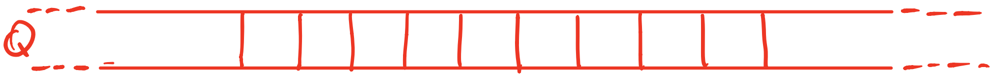
	f	left	right
a_1	5	/	/
a_2	9	/	/
a_3	12	/	/
a_4	13	/	/
a_5	16	/	/
a_6	45	/	/
b_1	14	a_1	a_2
b_2	25	a_3	a_4
...	...	...	...
...	...	...	...
...	...	...	...



	f	left	right
a_1	5	/	/
a_2	9	/	/
a_3	12	/	/
a_4	13	/	/
a_5	16	/	/
a_6	45	/	/
b_1	14	a_1	a_2
b_2	25	a_3	a_4
b_3	30	b_1	a_5
b_4	55	b_2	b_3
b_5	100	a_6	b_4

	f	left	right
a_1	5	/	/
a_2	9	/	/
a_3	12	/	/
a_4	13	/	/
a_5	16	/	/
a_6	45	/	/
b_1	14	a_1	a_2
b_2	25	a_3	a_4
b_3	30	b_1	a_5
b_4	55	b_2	b_3
b_5	100	a_6	b_4





- POICHE' VALE $f(b_1) \leq f(b_2) \leq \ldots \leq f(b_{n-1})$ (DA DIMOSTRARE!), AD OGNI ITERAZIONE PER TROVARE I DUE MINIMI x', x'' E' SUFFICIENTE ISPEZIONARE IN TEMPO $O(1)$ AL PIÙ I DUE ELEMENTI IN CIMA ALLE CODE Q_1 E Q_2.

- DUNQUE COMPLESSIVAMENTE L'ALGORITMO DI HUFFMAN CON PREPROCESSING IMPIEGA TEMPO $O(n)$.

- OVVIAMENTE IL PREPROCESSING POTRA' RICHIEDERE TEMPO $O(n \log n)$.

CORRETTEZZA

PER SEMPLICITÀ, SCRIVEREMO x AL POSTO DI $f(x)$

- E' SUFFICIENTE DIMOSTRARE CHE

$$b_1 \leq b_2 \leq \dots \leq b_{n-1}$$

- SIANO $Q_1^{(i)}$, $Q_2^{(i)}$, PER $i = 0, 1, 2, \dots, n-1$ LE CONFIGURAZIONI DELLE CODE Q_1 E Q_2 DOPO L'i-ESIMO PASSO

DUNQUE

$$Q_1^{(0)} := [\infty, a_n, \dots, a_1]$$

$$Q_2^{(0)} := [\ ]$$

- SIANO $x_i' \leq x_i''$ I DUE MINIMI DI $Q_1^{(i)} \cup Q_2^{(i)}$, PER $i = 0, 1, 2, \dots, n-1$

DUNQUE: $x_0' := a_1$

$x_1' := a_2$

- SIA $m_i = \min\left(Q_1^{(i)} \cup Q_2^{(i)} \right)$

DUNQUE $m_0 = a_1$

- DIMOSTRIAMO PER INDUZIONE SU $i = 0, 1, \ldots, n-2$

CHE: (1) $Q_2^{(i)}$ E' ORDINATA

(2) $\left(m_i = \right) x_i' \leq x_i'' \leq m_{i+1}$

- <u>CASO BASE $i=0$</u>: $Q_2^{(0)}$ E' ORDINATA

$\{x_0', x_0''\} = \{a_1, a_2\}$

PERTANTO:

POTREBBE ESSERE ∞

$$\left.\begin{array}{l} a_1 \leq a_2 \leq a_3 \\ a_1, a_2 \leq a_1 + a_2 = b_1 \end{array}\right\} \Longrightarrow a_1, a_2 \leq m_1$$

QUINDI: $x_0' \leq x_0'' \leq m_1$

- **PASSO INDUTTIVO:**

SIA $1 \leq i \leq n-2$ E SUPPONIAMO CHE

- $Q_2^{(i-1)}$ SIA ORDINATA
- $x'_{i-1} \leq x''_{i-1} \leq m_i$

POICHE' $m_i \leq x'_i, x''_i$,

$$b_i = x'_{i-1} + x''_{i-1} \leq x'_i + x''_i = b_{i+1}$$

QUINDI $Q_2^{(i)}$ E' ORDINATA

INOLTRE:

- $TOP(Q_1^{(i+1)}) \geq x''_i$
- SE $|Q_2^{(i+1)}| > 1$, BANALMENTE $TOP(Q_2^{(i+1)}) \geq x''_i$

  ALTRIMENTI, $TOP(Q_2^{(i+1)}) = b_{i+1} = x'_i + x''_i \geq x''_i$

IN OGNI CASO SI HA: $TOP(Q_2^{(i+1)}) \geq x_i''$

PERTANTO $m_{i+1} = \min\left(TOP(Q_1^{(i+1)}), TOP(Q_2^{(i+1)})\right) \geq x_i''$

E DUNQUE: $x_i', x_i'' \leq m_{i+1}$.