FUNZIONE DI HASH

U universo m n° di bucket

$$h: U \longrightarrow m$$

1) h sia calcolabile in tempo cost.
2) h occupi spazio costante
3) (Full randomness assumption)
Si possa estrarre un'fo random
a caso una f. di hash
(fissati U e m)

PEELING SEQUENCE

$G = (V, E)$ non orientato

Peeling sequence per G

e_1, x_1

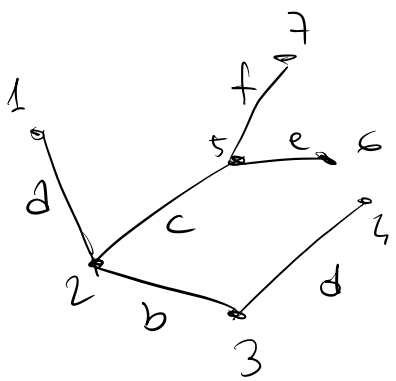
e_2, x_2

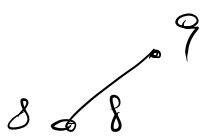
$\vdots$

e_m, x_m

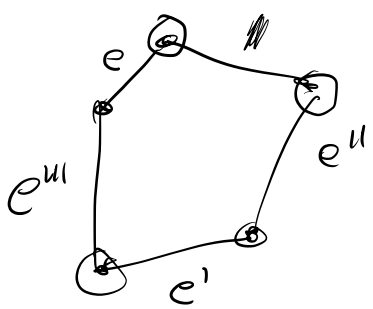
1) $e_1, \dots, e_m$ sono una perm. di E
2) $x_i \in e_i$
3) $x_i \notin e_1 \cup \dots \cup e_{i-1}$

$\uparrow E \quad \uparrow V$ ← hinge

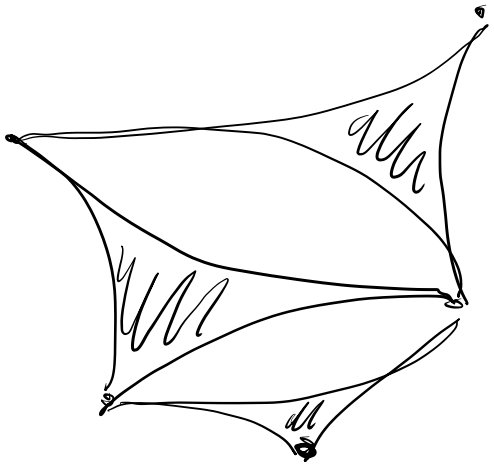




g, 8
a_1, 1
f, 7
e, 6
d, 4
c, 5
b, 2



3 - ipergrafo



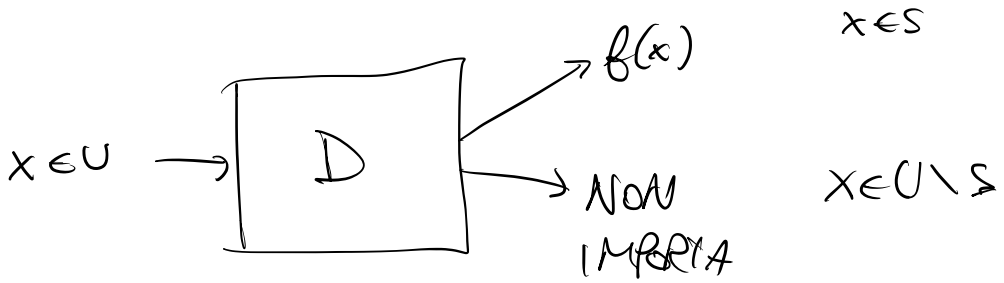
MEMORIZZAZIONE DI UN DIZIONARIO STATICO

$U \supseteq S$

$|S| = n \ll |U|$

S chiavi

$f : S \to 2^r$ PRESCRITTA



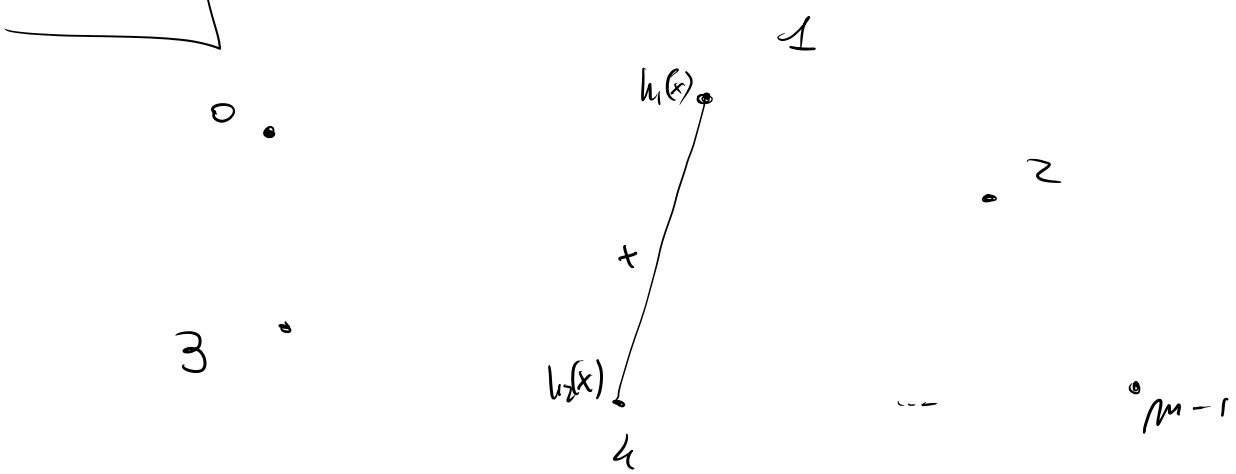
MWHC = Majewski - Worwald - Havas - Czech

m fissato intero

m lati
$n = |S|$
vertici

$h_1, h_2 : U \to m$

un lato per ogni $x \in S$

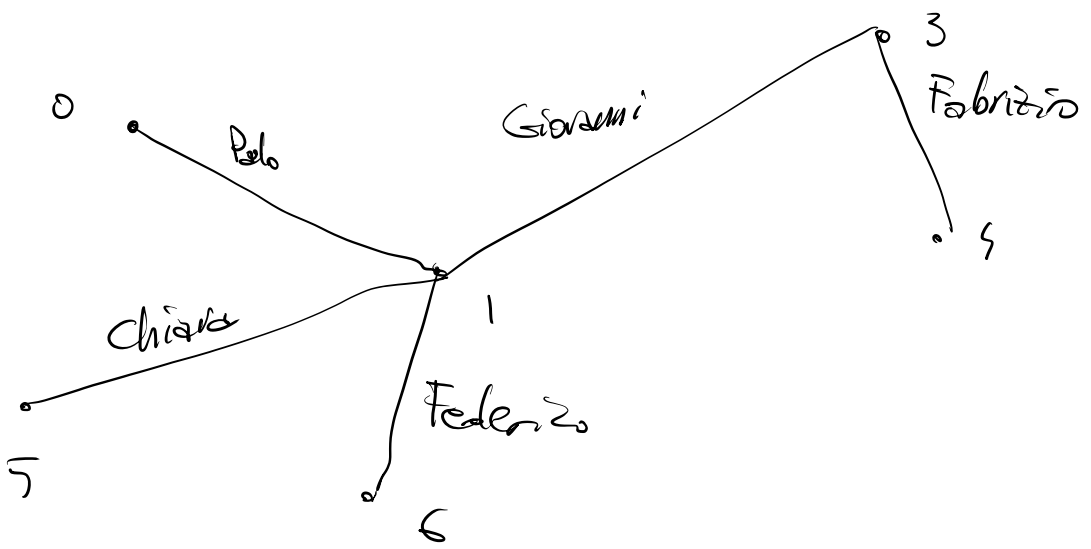


1) $\{h_1(x), h_2(x)\} = \{h_1(y), h_2(y)\}$

2) $h_1(x) = h_2(x)$

3) grafo ciclico

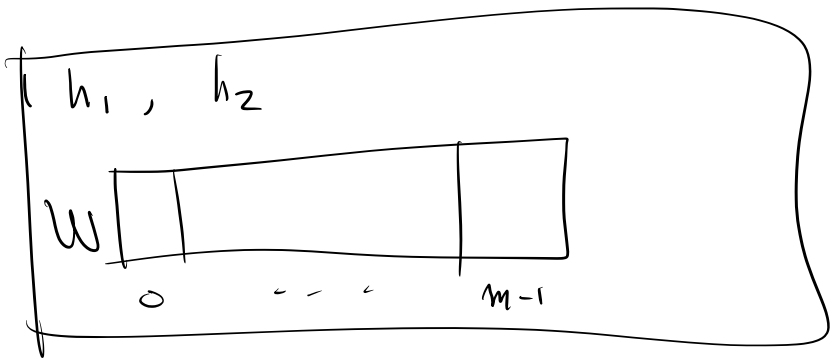
] No!



$$\begin{cases} (W_{h_1(x)} + W_{h_2(x)}) \bmod 2^r = f(x) & x \in S \\ \vdots \\ \vdots \end{cases}$$

m eq.

$W_0, W_1, \ldots, W_{m-1}$ variabili



m r bit

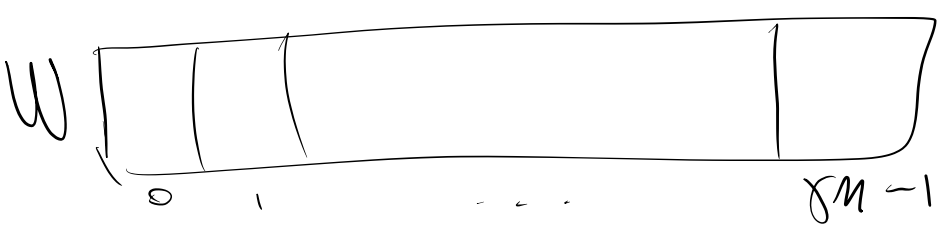
$f(\text{"Chiara"}) = W[h_1(\text{"Chiara"})] + W[h_2(\text{"Chiara"})] \bmod 2^r$

Teorema: Se $m > 2.09n$, la procedura di sel. di h_1, h_2 termina con prob. $\geq 1/2$ (In $E[\cdot]$ 2 tentativi)

$\boxed{\underbrace{2.09n}_{m} r \text{ bit}}$

3-iperarchi $\quad m > \underbrace{1.23}_{\gamma} n$

$\boxed{\gamma n r \text{ bit}}$



VARIANTE COMPRESSA

1) memorizzano solo le $\leq n$ entry di W non nulle (nr bit)

2) memorizza l'array di bit che dice dove sono e su quale rank (γn bit) ($o(n)$ bit)

$r+\gamma$ bit per chiave	$nr + \gamma n + o(n)$	COMPRESSA
$r\gamma$ bit per chiave	γnr	NON COMPRESSA

$$r + \gamma < r\gamma$$

$$r(\gamma - 1) > \gamma$$

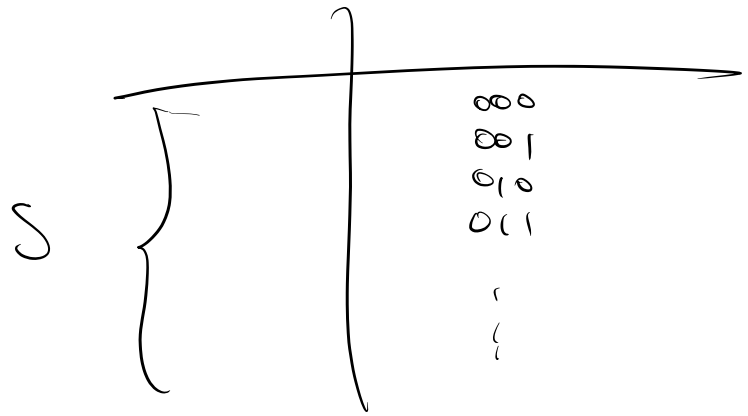
$$r > \frac{\gamma}{\gamma - 1}$$

$$r > 5$$

HASH MINIMALE PERFETTO

$$h : U \to m$$

- è perfetto per $S \subseteq U$ se h è iniettiva su S
- se $m = |S|$ minimale



$n \lg n + pn + o(n)$ bit

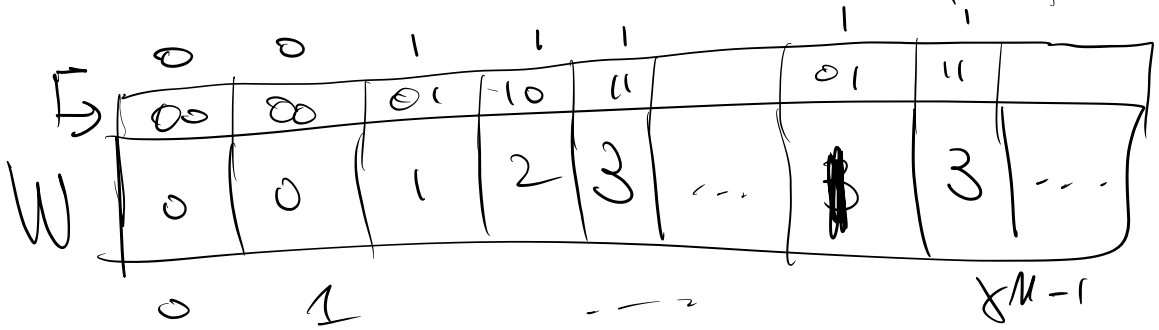
$$\begin{cases} (W_{h_0(x)} + W_{h_1(x)} + W_{h_2(x)}) \bmod 3 = 1 \\ \vdots \end{cases} \qquad r = 2$$

$$\to \boxed{H(x)} = h(x)_{(W[h_0(x)] + W[h_1(x)] + W[h_2(x)]) \bmod 3}$$

iniettiva su $\{0,1,\dots,m-1\}$

$\Rightarrow$ perfetta ma non minimale

$$2m\gamma + o(n) + \cancel{o(n)}$$



$W_i = 0$ sse i non è hinge

$(2\gamma)m$ bit

2.46 bit/elemento

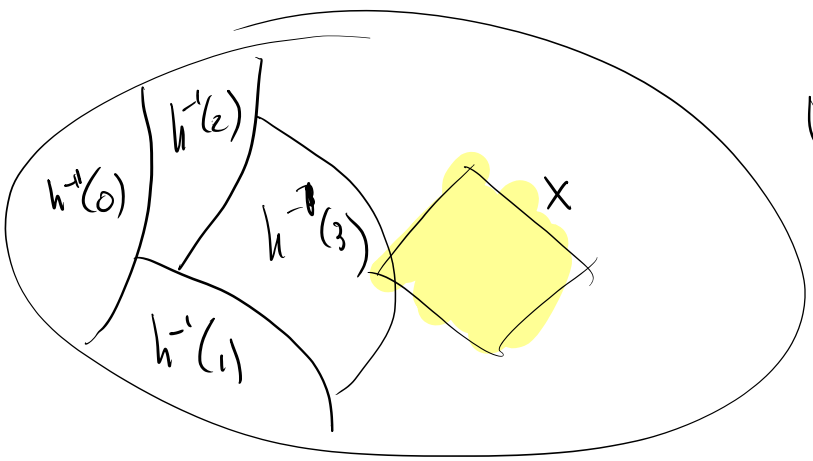
INFORMATION-THEORETICAL LOWER BOUND

Quante sono le funzioni di hash

$$h: U \to n$$

minimali perfette per $X \subseteq U$, $|X| = n$.

$$h: U \to n$$



U

$|X| = n$

h separa X se ogni elemento di X viene messo in parte diversa

$H_U(n)$ = minimo numero di funz. $h: U \to n$ che separa ogni $X \subseteq U$ $|X| = n$

$$Vol(h) = \#\text{ insiemi che } h \text{ separa} \quad \times \quad (|X| = n)$$

$V_{MAX} = Vol(h)$ massimizzato se h ha le fibre $\approx \frac{|U|}{n}$

$$H_U(n) \geq \frac{\binom{|U|}{n}}{V_{MAX}}$$

$$V_{MAX} \approx \left(\frac{|U|}{n}\right)^n$$

$$H_U(n) \geq \frac{\binom{|U|}{n}}{\left(\frac{|U|}{n}\right)^n}$$

$n \leq \sqrt{u}$ (2ε)

.....

$$\ln \quad H_U(n) \geq n + O(\ln n)$$

$$\log H_U(n) \overset{U}{=} \frac{\ln H_U(n)}{\ln 2} \Rightarrow 1.44n + O(\log n)$$