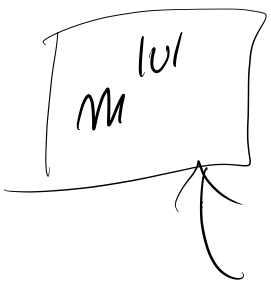
FUNZIONI DI HASH

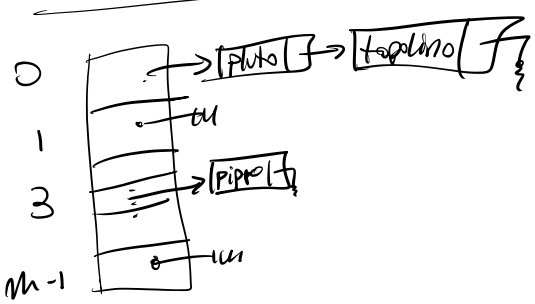
- Universo U
- $m \in \mathbb{N}$ numero di bucket

$h : U \longrightarrow m$



funzioni di hash per U con m bucket

$\mathcal{H}_{U,m}$



$S \subseteq U$

$\Sigma = \{a, b, c, d\}$ $\quad 0\ 1\ 2\ 3$ $\quad \Sigma^*$

$m \in \mathbb{N}$

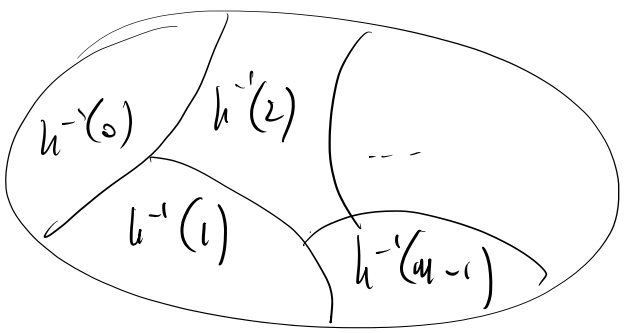
$h : \Sigma^* \longrightarrow m$

$h(\text{"pippo"}) = 3$

$h(\text{"topolino"}) = 0$

$h(\text{"pluto"}) = 0$

Σ^*



ASSUNZIONI

1) possiamo estrarre uniformemente a caso $h \in \mathcal{H}_{U,m}$
[FULL RANDOMNESS ASSUMPTION]

2) h sia calcolabile in tempo e spazio costante e occupi spazio costante

$\Sigma^{\leq k}$

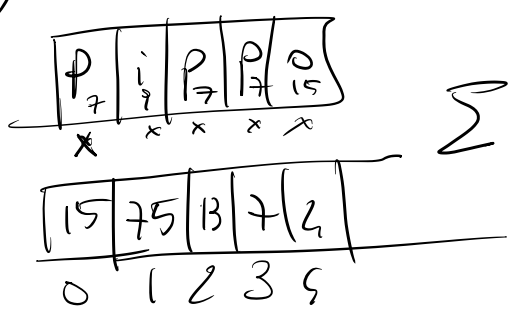
$h : U = \Sigma^{\leq k} \longrightarrow m$

M^k → 

Valori pseudocasuali in $\{0, \dots, m-1\}$

$\Sigma = \{a_0, b_1, c_2, d_3, \dots\}$

$h("pippo")$



$\Sigma \longrightarrow$ modulo

SEQUENZA DI PEELING DI UN GRAFO

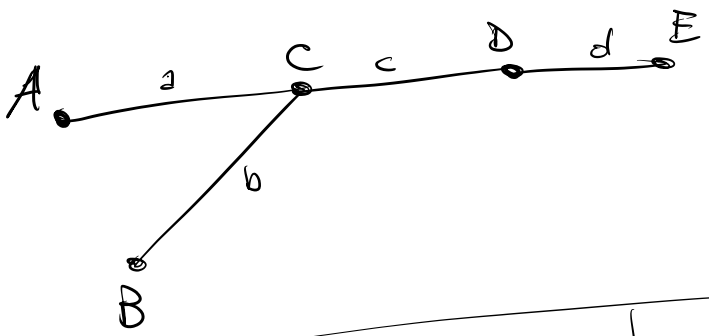
$G = (V, E)$ non orientato

Una sequenza di peeling

tutti i lati di E $\left\{\begin{array}{ll} e_0, & x_0 \\ e_1, & x_1 \\ e_2, & x_2 \\ \vdots & \\ e_{m-1}, & x_{m-1} \end{array}\right.$ hinge — verifica

$x_i \in e_i$

t.c. $\forall i$ x_i non è usato compreso nei lati che precedono i



G

$a = \{A, C\}$, A
$b = \{B, C\}$, B
$c = \{C, D\}$, D
$d = \{D, E\}$, E

$a = \{A, C\}$, A
$b = \{B, C\}$, C No!

Teorema: G ammette una sequenza di peeling se e solo se è aciclico.

Dim: $\Rightarrow$

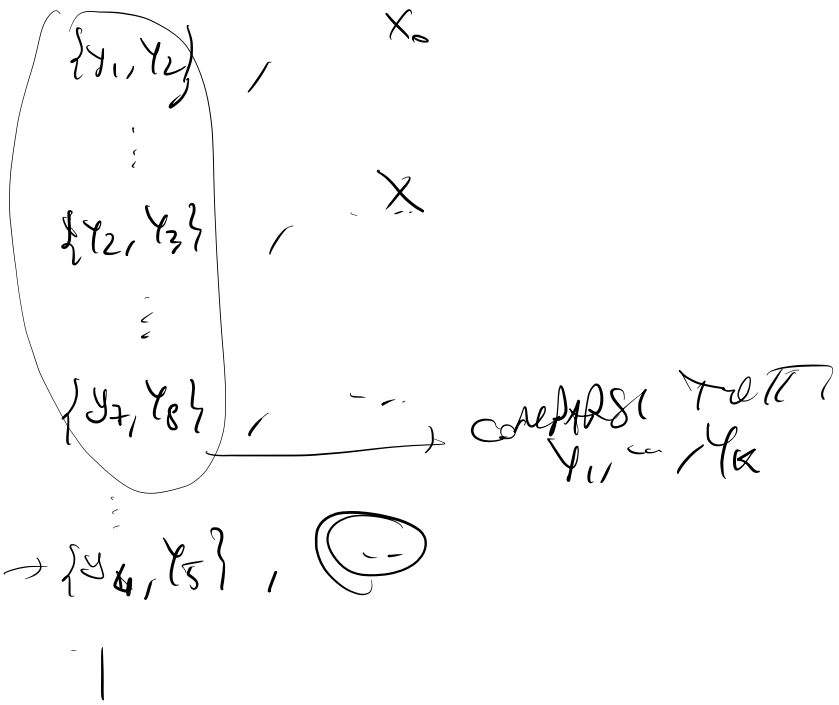
$\left.\begin{array}{l} e_0, x_0 \\ \vdots \\ e_{n-1}, x_{n-1} \end{array}\right\}$ seq. di peeling

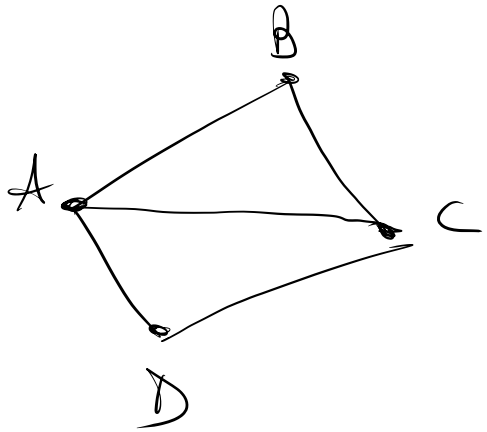
se c'è un ciclo $y_1, y_2, \dots, y_k$

$\{y_1, y_2\}$ $\{y_2, y_3\}$, $\dots$, $\{y_k, y_1\}$

e_{i_1} e_{i_2} e_{i_k}

sia i l'indice massimo





$\{A,B\}$, A

$\{A,D\}$, D

$\{D,C\}$, C

$\{B,C\}$

$|E|$ Per induzione su $|E|$.
(MESSA) . □

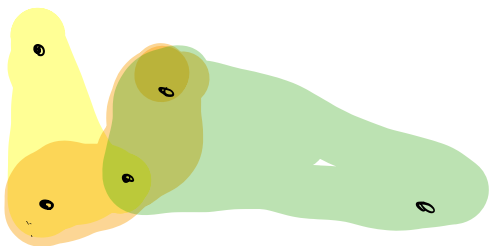
IPERGRAFI

GENERALIZZAZIONE dei grafi non orientati

r- IPERGRAFO

$G = (V, E)$ $\quad E \subseteq \binom{V}{r}$

vertici, iperlati



SEQ. DI PEELING

MEMORIZZARE FUNZIONI STATICHE

- U universo
- $X \subseteq U$
- $r \in \mathbb{N}$

sottoinsieme finito dell'universo

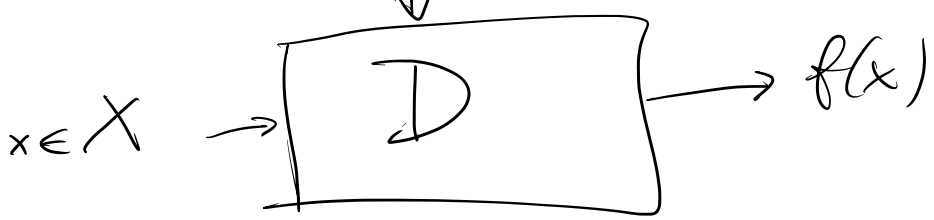
$$f : X \to 2^r$$

VOGLIAMO MEMORIZZARE f

$U = \Sigma^*$ $\quad \Sigma = ASCII$ $\quad r = 5$

$X \subseteq U$	$f(x)$	
Paolo Boldi	00111	7
Anna Zoppi	10100	20
Giovanni Gelli	10111	23

← VOGLIO MEM. QUESTO



TECNICA MWHC (Majewski, Worwald, Havas, Czech)

Fisso un $m \in \mathbb{N}$ e scelgo

$h_1, h_2 : U \longrightarrow m$

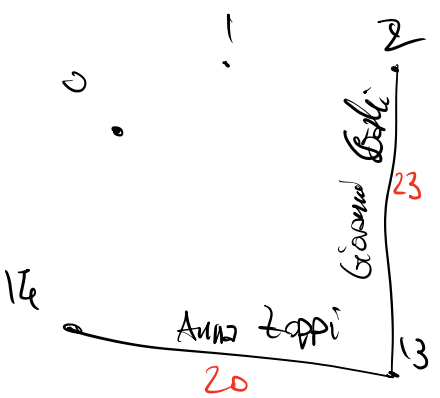
Costruisco un grafo

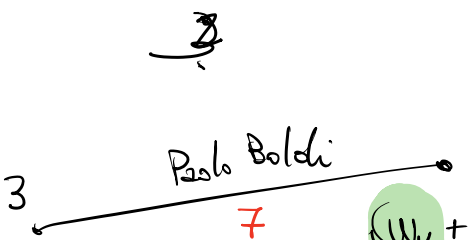
- vertici $\{0, 1, \dots, m-1\}$
- lati $x \in X$ $\{h_1(x), h_2(x)\}$

$m = 17$

$X \subseteq U$	$f(x)$		h_1	h_2
Paolo Boldi	00111	7	3	7
Anna Zoppi	10100	20	14	13
Giovanni Galli	10111	23	13	2

$w_0, w_1, \dots, w_{m-1}$





$w_{14} = 20$ $w_7 = 7$

$w_2 = 23$

$$\begin{cases} (w_4 + w_{13}) \bmod 2^5 = 20 \\ (w_2 + w_{13}) \bmod 2^5 = 23 \\ (w_3 + w_7) \bmod 2^5 = 7 \end{cases}$$

1) $\forall x \in X. \quad h_1(x) \neq h_2(x)$

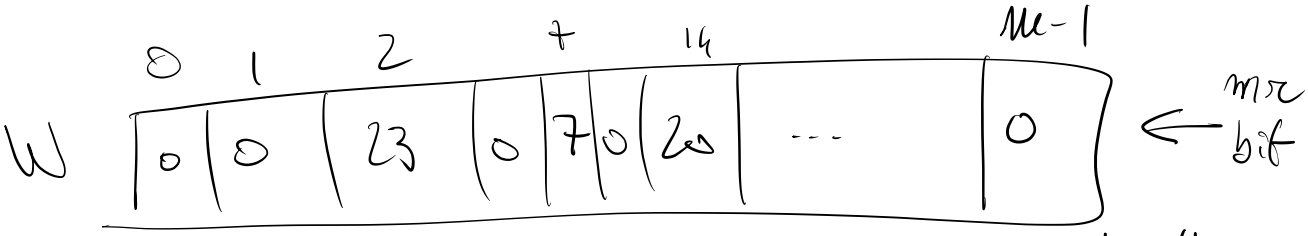
2) $\forall x, y \in X \quad x \neq y$

$\{h_1(x), h_2(x)\} \neq \{h_1(y), h_2(y)\}$

→ 3) $\in$ deve essere aciclico

EQUAZIONE $\forall x \in X$

→ $\left(W_{h_1(x)} + W_{h_2(x)}\right) \bmod 2^r = f(x)$



$f(\text{"Paolo Boldi"}) \longrightarrow$ $h_1(\text{"Paolo Boldi"})$, $h_2(\text{"Paolo Boldi"})$

$f(\text{"Mario Rossi"})$

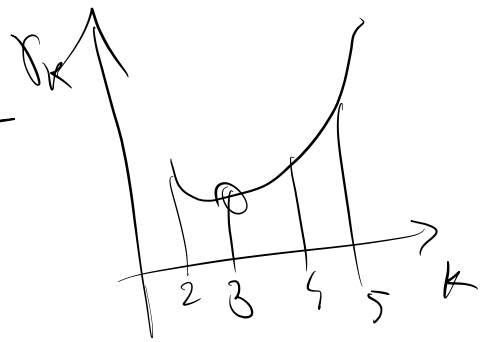
Teorema: Se $m > 2.09n$ $(n = |X|)$, il grafo è quasi sempre aciclico. Il numero atteso di tentativi è ≈ 2.

$m > 2.09n$ #BIT $2.09nr$

+ bit necessari per memorizzare h_1, h_2

Teorema: Per ogni k, esiste una costante γ_k t.c. se $m > \gamma_k n$ allora l'ipergrafo quasi sempre ammette una funzione di peeling

k	γ_k
2	2.09
3	1.23
4	> 1.23
5	$>$



CORDATA

BIT

$1.23nr$

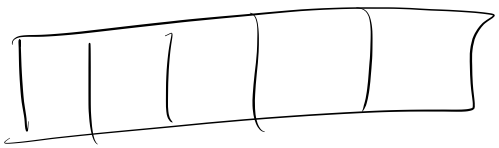
$f: X \to 2^r$

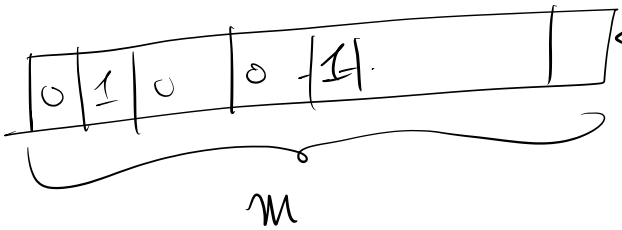
$(2^r)^{|X|} = 2^{rn}$

INF. TR. L.B.

rn

In W (che ha $m = 1.23n$ elementi) solo $\leq n$ sono non zero.

$n\pi$ bit $\rightarrow$ $\tilde{W}$  ← elementi non nulli di W

m bit $\rightarrow$ b  ← posizioni degli el. non-nulli in W

$$W[i] = \begin{cases} 0 & \text{se } b_i = 0 \\ \tilde{W}[\underline{\text{rank}_b}(i)] & \text{se } b_i = 1 \end{cases}$$

$n\pi + m$	bit	$1.23n\pi$
$n\pi + 1.23m$	bit	
$(\pi + 1.23)n$	bit	NON COMPRESSA

COMPRESSA

$$(\pi + 1.23)m < 1.23n\pi$$

$$\pi + 1.23 < 1.23\pi$$

$$\pi > 5$$

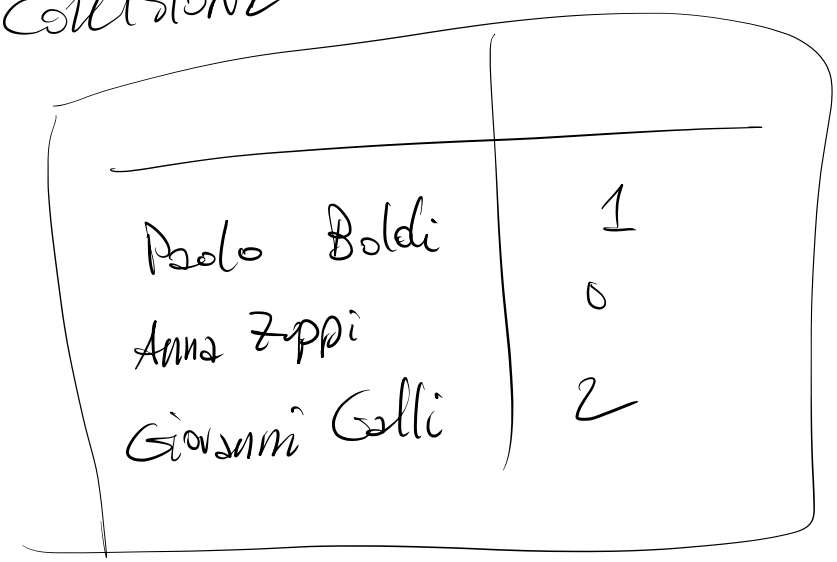
FUNZIONE DI HASH MINIMALE PERFETTA

$X \subseteq U$

$|X| = n$

n° di bucket = n° di dati

PRIVO DI COLLISIONE



MWHC