

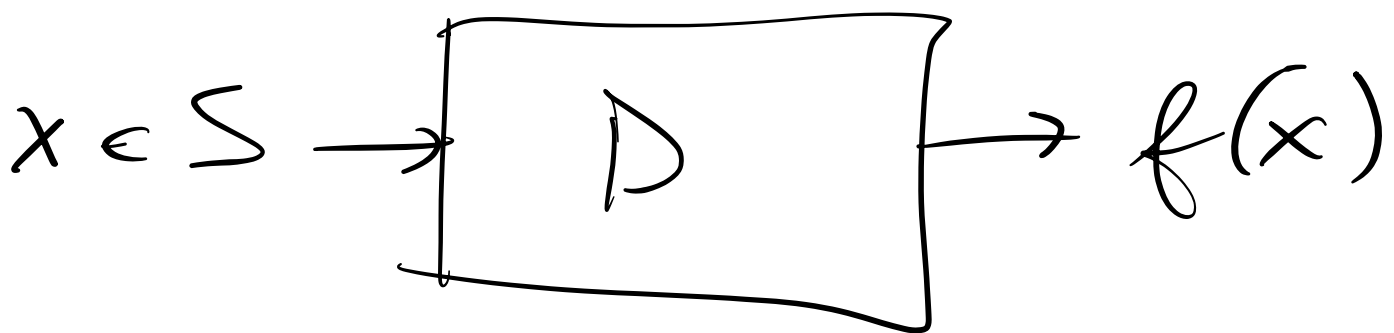
FUNZIONI STATICHE

Dato U e un $n > 0$,
ho una funzione

$$f: S \rightarrow 2^n$$

con $S \subseteq U$.

Voglio rappresentare f .



f	S	INDIRIZZO
	"Paolo Baldi"	via Giovanni Tonti, 7
	"Pippo Berti"	...
	"Cina"	

Giuliana
Lassi
...

$$U = \sum_{\leq 100}$$

$$\Sigma = \text{ASCII}$$

TECNICA

MWHC

$$S \subseteq U$$

$$f: S \rightarrow 2^{\mathbb{N}}$$

$$|S| = m$$

Fisso

$$m \in \mathbb{N}$$

e due

funzioni

$$h_0, h_1: U \rightarrow \mathbb{N}$$

Costruisco

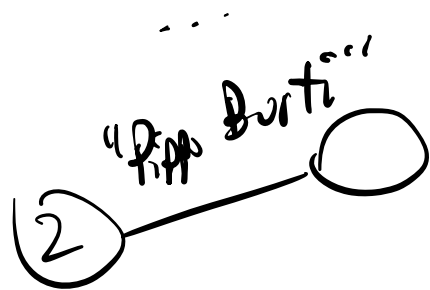
un

grafo

①

①

①



③

$s \in S$

$\rightarrow$ $\begin{matrix} \text{bto} \\ \text{M} \\ \text{bti} \end{matrix}$

$h_0(s), h_1(s)$

1) $h_0(x) \neq h_1(x)$

$x \in S$

2) $\{h_0(x), h_1(x)\} \neq$

$\forall x, y \in S$

$\{h_0(y), h_1(y)\}$

$x \neq y$

3) $\subseteq$ sia aciclico



Vedo $\hookrightarrow$ come un
insieme di eqvt.
 n equazioni

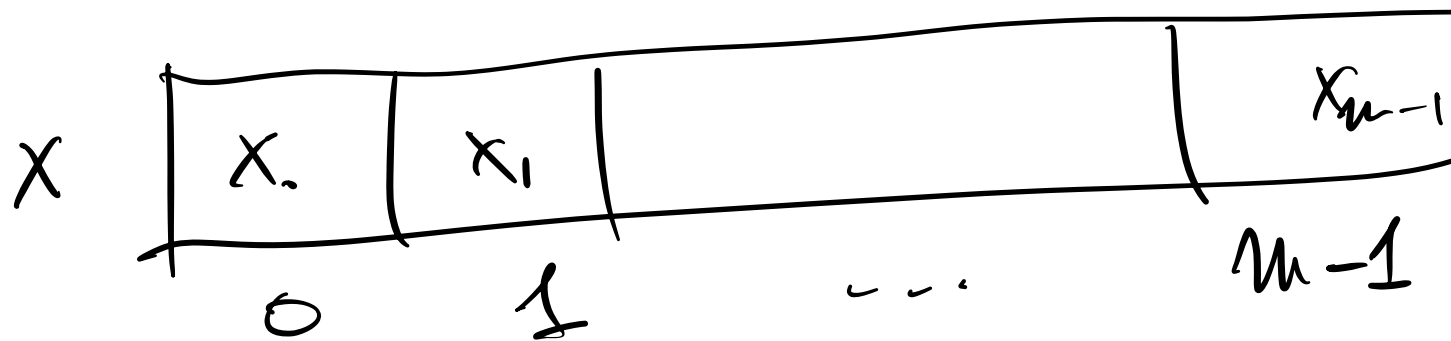
$$\left(X_{h_0}(\text{"Bolo Boldi"}) + X_{h_1}(\text{"Bolo Boldi"}) \right)$$

$$\text{mod } 2^{\Omega} = \text{" via Giovanni Tonti, 7"}$$

$$\left\{ \begin{array}{l} (X_{h_0(s)} + X_{h_1(s)}) \text{mod } 2^{\Omega} = f(s) \\ \forall s \in S \end{array} \right.$$

M equazioni distorte
con m incognite $(x_0, \dots, x_{m-1})$

ACICLICITÀ $\Rightarrow$ SOLUBILITÀ



$\uparrow$
SOLUZIONE DEL
SISTEMA

$m \pi$ bit

$$\left(x_{h_0(s)} + x_{h_1(s)} \right) \bmod 2^\pi = f(s)$$

Teorema: Se $m > 2.09n$

il problema è quasi sempre

decidico (e bastano
2 tentativi)

$$M = 2 \cdot \Theta(n)$$

$2 \cdot \Theta(n) \text{ bit}$

PELABILITA' DI UN IPERGRAFO

Un ipergrafo è pelabile
se esiste un ordinamento

e_0

e_1

$\vdots$

HINGE

x_0

x_1

$\vdots$

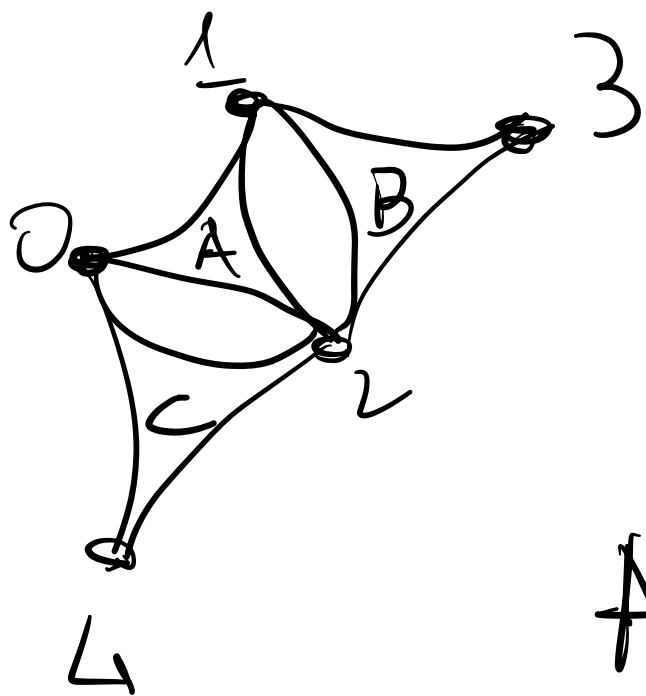
x

$x_i \in e_i$

E_{k-1}
↑
iperlati

Λ_{k-1}
↑
vertici

t.c. $x_i \notin E_0 \cup \dots \cup E_{i-1}$



A

0

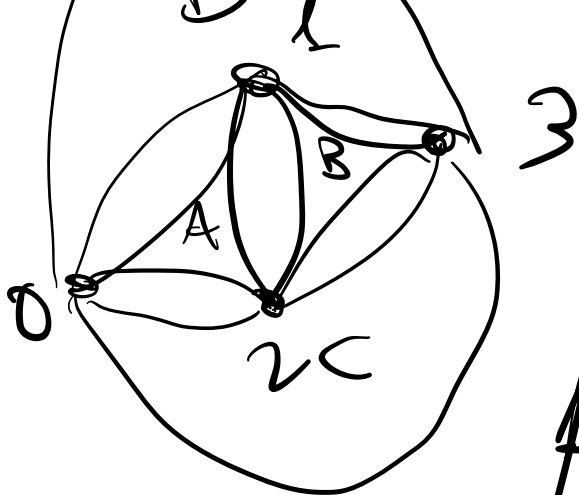
$\subset$

4

B

3

D.



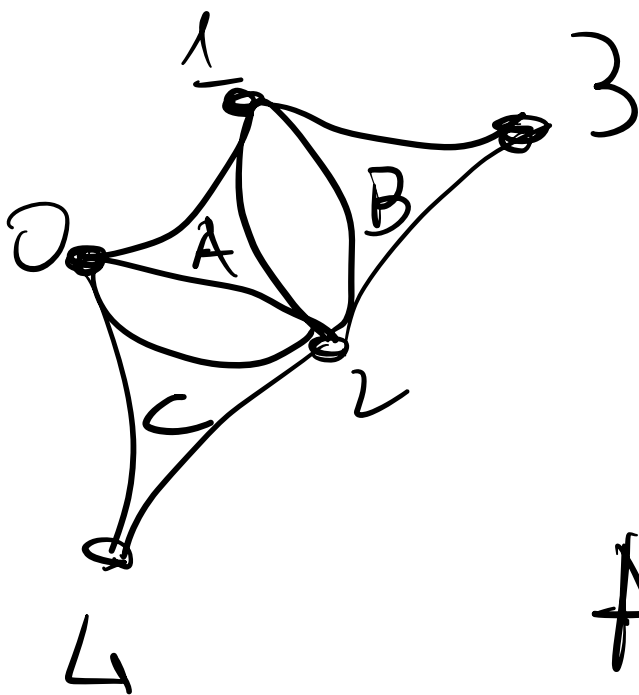
A

0

B

3

D



A

0

C

4

B

3

mod 100
↓

$$= 25$$

$$= 37$$

$$= 12$$

A

$$x_0 + x_1 + x_2$$

C

$$x_0 + x_2 + x_4$$

$$x_1 + x_2 + x_3$$

B

25	0	0	12	12
----	---	---	----	----

1	0	0	1	1
x_0	x_1	x_2	x_3	x_4

Teorema: Per ogni
 $k \geq 2$, esiste γ_k
 t.c. $\forall m \geq \gamma_k^n$
 la costruzione del
 k -ipergrafo è
 possibile e soddisfa
 le altre due cond.
 quasi certamente.

$\gamma_2 = 2.09$

$$\boxed{\gamma_3 = 1.23}$$

$$\forall k \geq 4 \quad \gamma_k > \gamma_3$$

INFORMATION
THE OPTIC LOWER
BOUND

$$f: S \rightarrow 2^{\mathbb{N}}$$

Per S fissato di
cardinalità n
sono $(2^{\mathbb{N}})^n$ funzioni.

$$Z_n = \log 2^{2^n} = 2^n$$

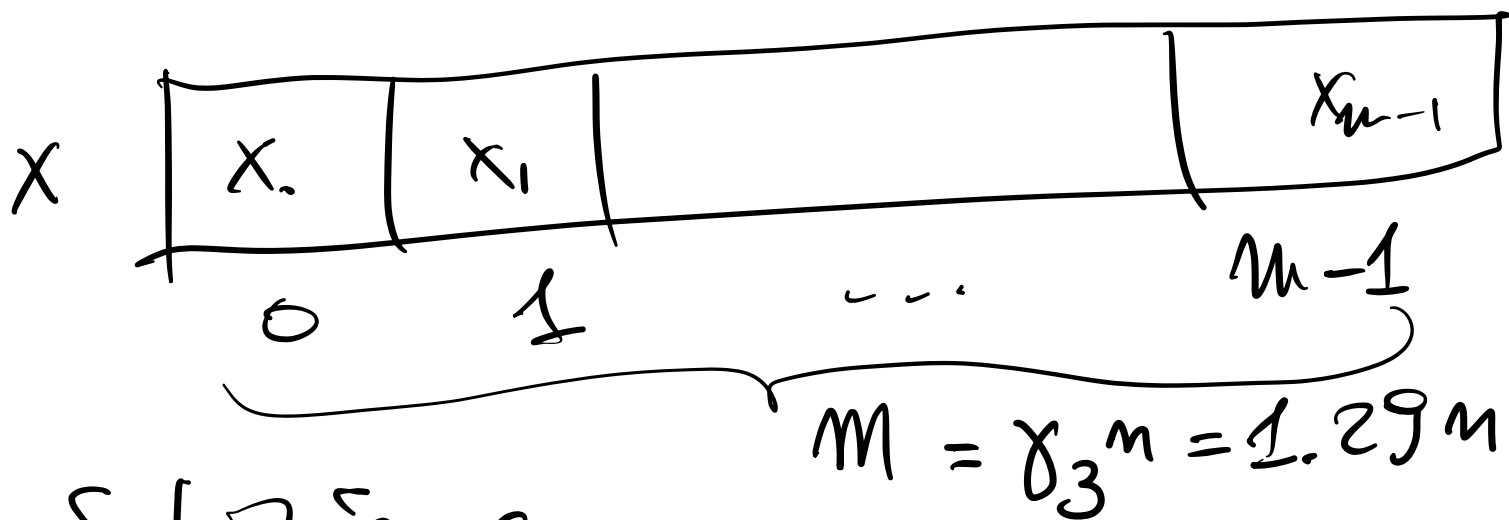
La nostra struttura
(con $k=3$) usa

$$D_n = \underbrace{\chi_3^n}_m R = \text{bit}$$

$$= 1.23 n R$$

COMPACTA

COMPRESSIONE DELL'ARRAY



Soluzione

$$m \text{ equazioni} \left\{ \begin{array}{l} (x_{h_0(s)} + x_{h_1(s)} + x_{h_2(s)}) = f(s) \\ \text{mod } 2^r \\ \forall s \in S \end{array} \right.$$

Nell'array ci sono $\leq m$ valori $x_i \neq 0$

Invece di X memorizziamo solo gli $\leq m$ elementi

non nulli e un
array di bit con
le loro posizioni.

$$n\pi + m + o(m) =$$

$\underbrace{\hspace{10em}}$
 ELEMENTI
NON
NULLI

$\underbrace{\hspace{10em}}$
 ARRAY
DI
BIT

$\underbrace{\hspace{10em}}$
 RANK/
SELECT
SUCCINTA

$$= m\pi + \gamma_3^m + o(\gamma_3^m) =$$

$$= m\pi + 1.23 m + o(m)$$

$$Z_m = m\pi$$

$\gamma \triangleq \gamma_3$

QUANDO CONVIENE?

$$\underbrace{n\pi + \gamma n + o(n)}_{\text{STR. COMPRESSA}} < \underbrace{n\gamma\pi}_{\substack{\text{STR.} \\ \text{NON} \\ \text{COMP.}}}$$

$$n\pi + \gamma n < n\gamma\pi$$

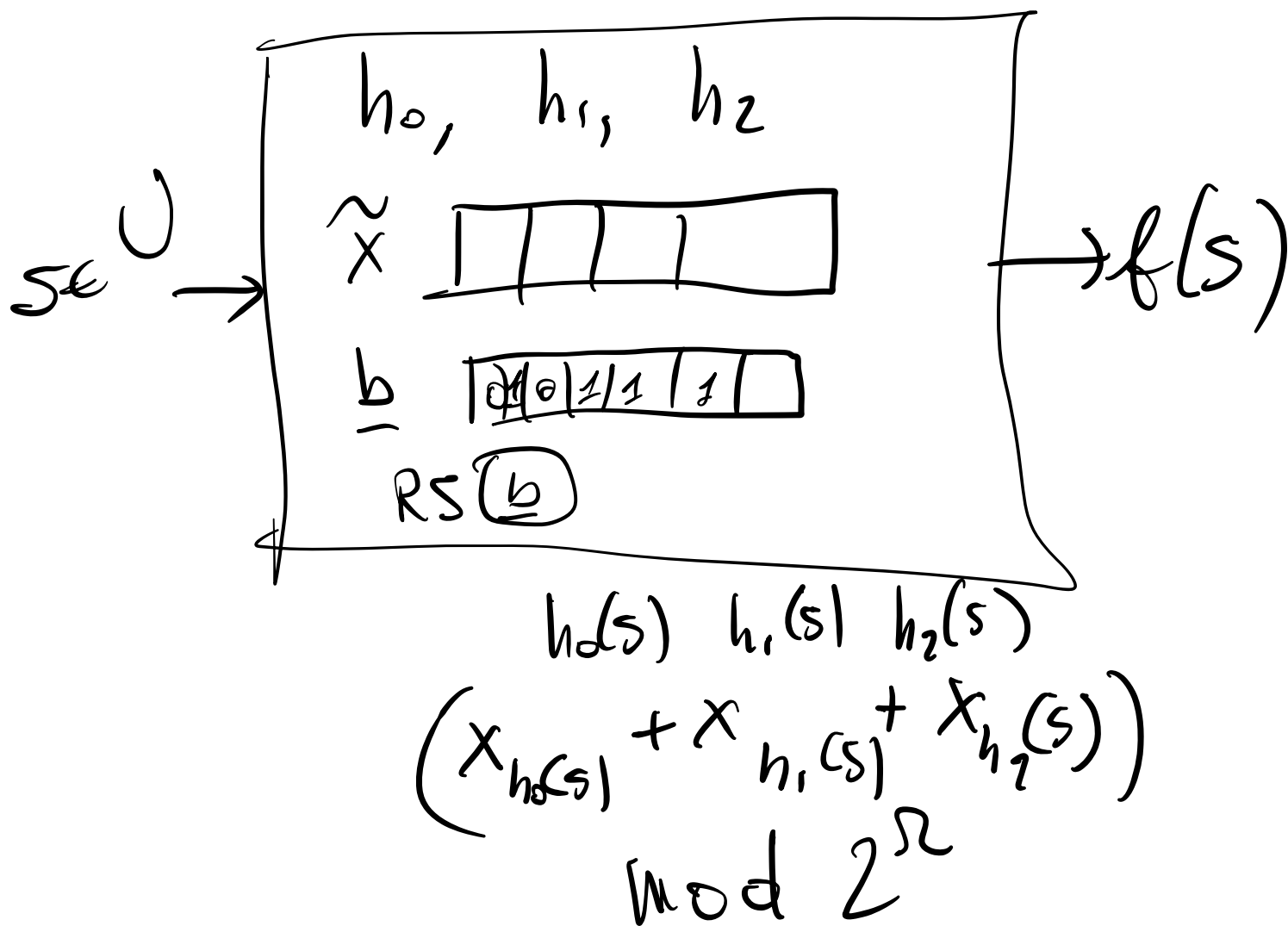
$$\pi + \gamma < \gamma\pi$$

$$(1 - \gamma)\pi + \gamma < 0$$

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

$$\pi > 5.$$

USO DELLA STRUTTURA



HASH MINIMALI PERFETTI

Sie $S \subseteq U$ fissi e
 $m \in \mathbb{N}$ fissato.

Una funzione di hash

$$h: U \rightarrow m$$

si dice

1) perfetta per S se
 $\forall x, y \in S \quad x \neq y \Rightarrow h(x) \neq h(y)$

$$\Rightarrow |S| \leq m$$

2) minimale perfetta se perfetta
e $|S| = m$. MPH

$$S \subseteq U$$

$$|S| = m$$

$h: U \rightarrow m$
minimale perfetta
per S

$A[h(s)]$

1° modo: fissa h A
PRIORI E USA MWHC

s	$h(s)$
"Paolo Boldi"	0
⋮	⋮
"Rosso Rosso"	$n-1$
	$\underbrace{\hspace{10em}}_{\pi = \log n}$

← MWHC

$$n\pi + \gamma n + o(n) \Rightarrow$$

$$\boxed{n \log n} + \gamma n + o(n)$$

OPMPH ← order Preserving
Minimality

Perfe VT
Hash

$$Z_n = \log n! \approx n \log n$$

2° CODD: USA UNA
VARIANTE DI
MWMC

Z_{MPH}

$< Z_{OPMPH}$

Z_{MPH}

$$S \subseteq U$$

MVHC

$$\left\{ \begin{array}{l} X_{h_0}(s) + X_{h_1}(s) + X_{h_2}(s) = \boxed{???} \\ \forall s \in S \end{array} \right.$$

$$X_{h_0}(s) + X_{h_1}(s) + X_{h_2}(s)$$

HINGE

$$X_{h_k}(s)$$

$$k=0,1,2$$

$$\left\{ \begin{array}{l} X_{h_0}(s) + X_{h_1}(s) + X_{h_2}(s) = h_k(s) \\ \forall s \in S \end{array} \right.$$

K H W G E

$\sim h$

PH $\nearrow$ $V_A \sigma R$

$0, \dots, m-1$

x	x_0	x_1	$\dots$	x_{m-1}
-----	-------	-------	---------	-----------

$\sim h$ SPA $z=0$ Occupatio

$n\pi + \chi n + o(n) =$

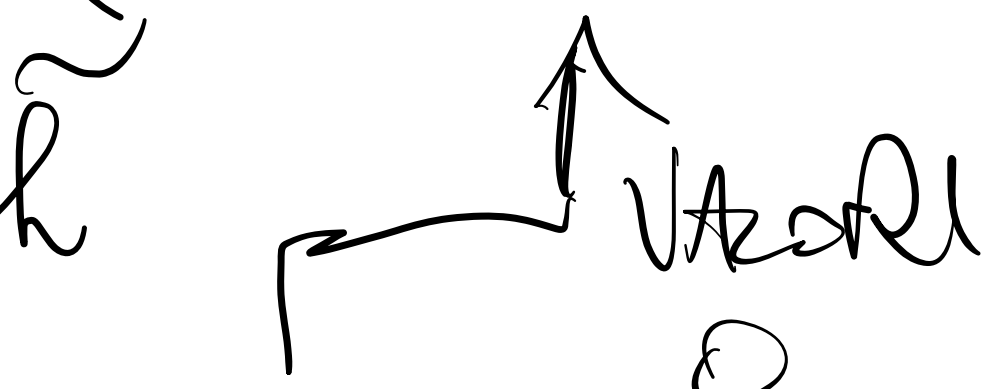
$$= n \log m + \gamma n + o(n)$$

$$= n \log(\gamma n) + \gamma n + o(n)$$

$$= n \log n + n \log \gamma + \gamma n + o(n)$$

$$\left. \begin{array}{l} X_{h_0(s)} + X_{h_1(s)} + X_{h_2(s)} \end{array} \right\} \text{HINGE} = \boxed{k}$$

$\forall s \in S$



$m-1$

PH

X

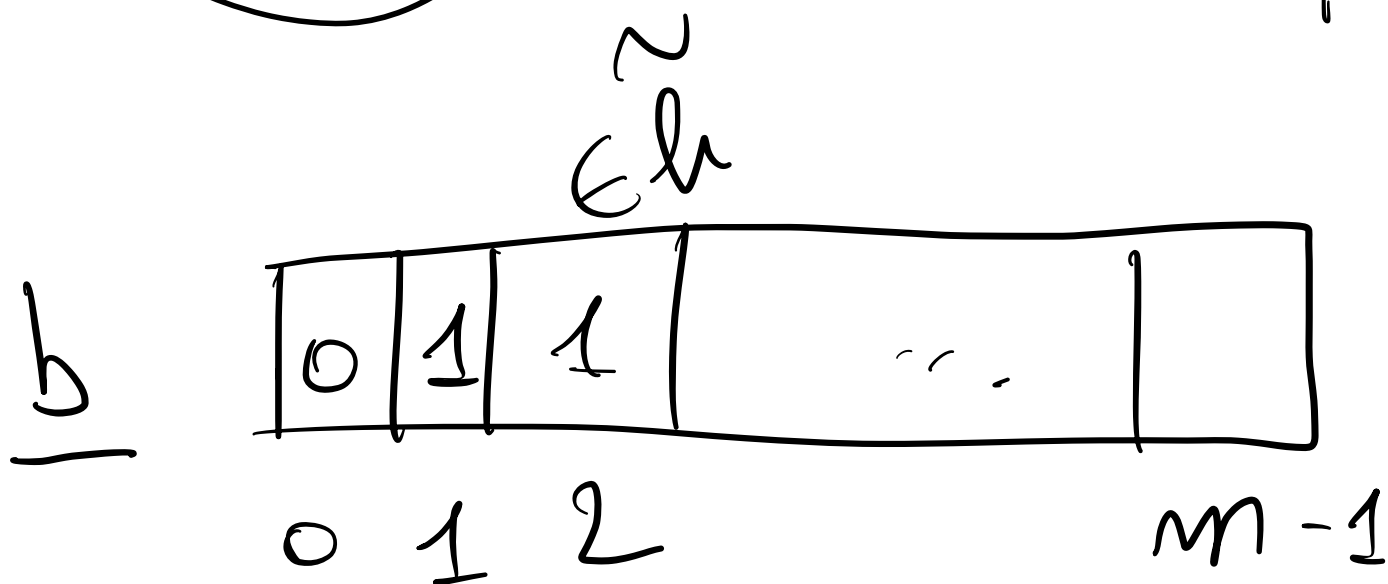
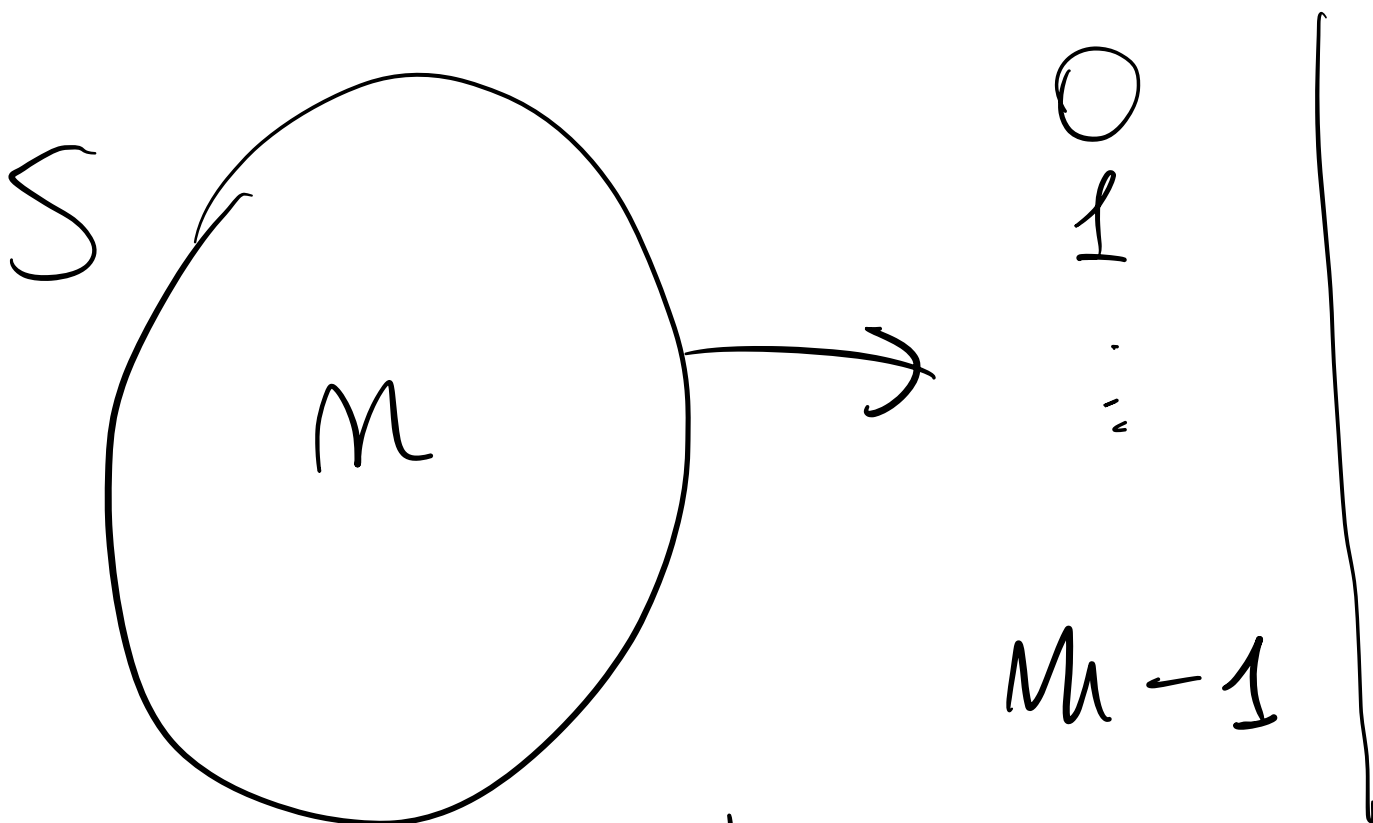
00	00	00	01	10	11
----	----	----	----	----	----

$$\tilde{h}(s) \triangleq$$

$$h(x_{h_0}(s) + x_{h_1}(s) + x_{h_2}(s)) \bmod 2^2$$

$$\begin{aligned} nR + \gamma n + o(n) &= \\ &= 2n + \gamma n + o(n) \end{aligned}$$

$$\tilde{h} : S \rightarrow M$$



m bit, n 1

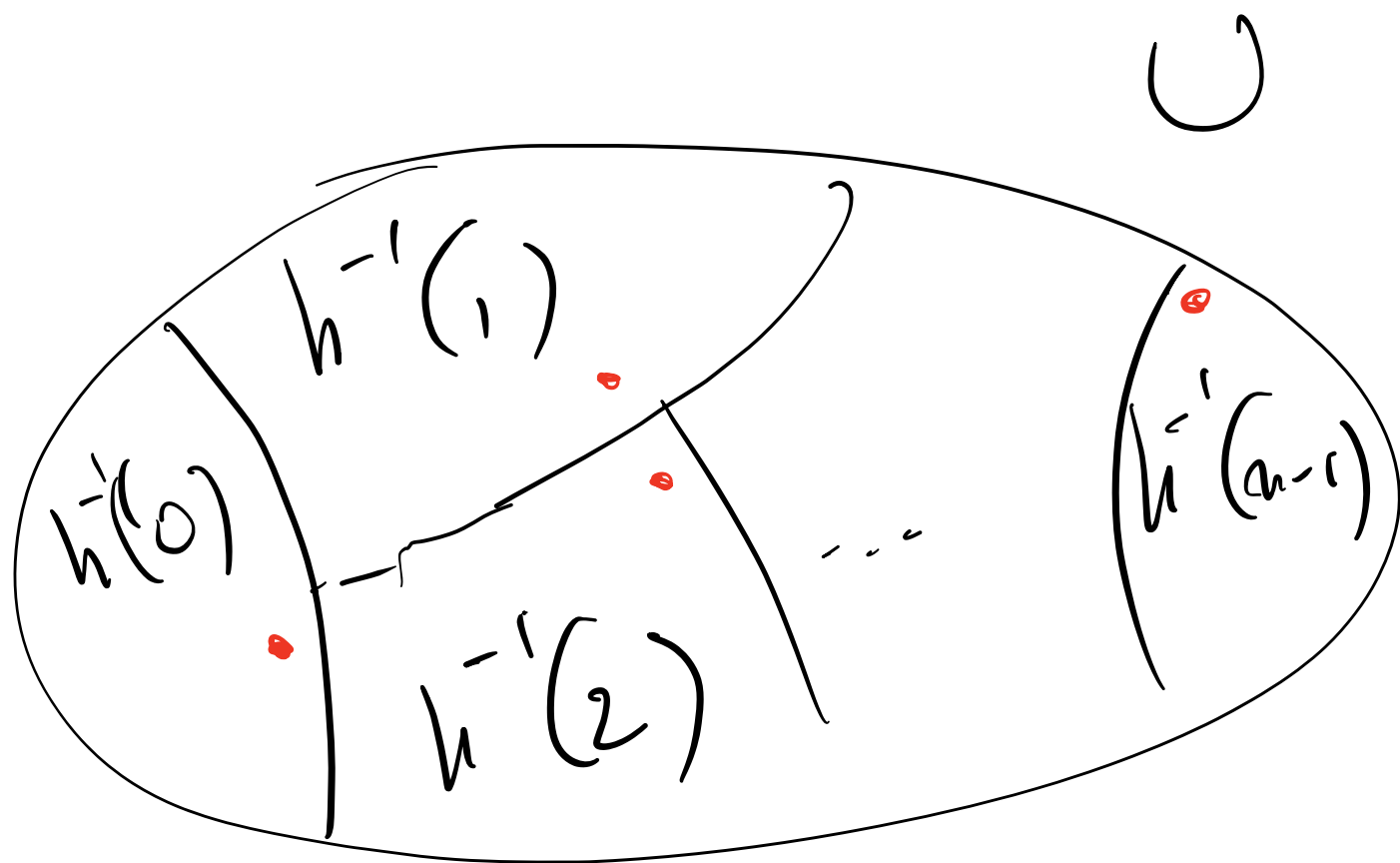
$$\underbrace{2n + \gamma n + o(n)}_{\sim n} + \underbrace{\gamma n}_{\text{red outline}} + o(n)$$

INFORMATION
THEORETICAL
LOWER BOUND

Quante sono le
MPH con m
...

Chiavi nell'universo
 U ?

$$h: U \rightarrow n$$



Per ogni S fatto scegliendo
 esattamente un $x \in h^{-1}(k)$

per ogni $k=0, \dots, n-1$,
 h è minimale perfetto
per S .

Si dice che h
SEPARA S .

$$h = \{h_0, h_1, h_2, \dots, h_{t-1}\}$$

h è un n -sistema

se $\forall S \subseteq U \quad |S|=n$
 $\exists i=0, \dots, t-1$

t.c. h_i separa S

Quanto è grande il
più piccolo n -sistema?

Chiediamo $H_U(n)$
la cardinalità del
più piccolo n -sistema
$$Z_n = \log H_U(n)$$

$$\binom{U}{n}$$

$$H_U(n) \geq \frac{1}{\sqrt{n}}$$

Ogni funzione di hash
 separa al massimo
 $\sqrt{n}$ insiemi di
 cardinalità n .

$$\sqrt{n} \left(\frac{1}{n} \right)^n$$

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



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

$$\log H_U(n) = \frac{\ln H_U(n)}{\ln 2} =$$

$$\geq \frac{n}{\ln 2} + O(\ln n)$$

$\ln 2$

$$Z_n \approx \underline{1.44 n} + O(\ln n)$$

$$\underbrace{2n + \gamma n + o(n)}_{\sim n} + \gamma n + o(n)$$

$$3.23n \rightsquigarrow 2.46n$$