Randomized Algorithms

南京大学

尹一通

Definition:

Events $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ are **mutually independent** if for any subset $I \subseteq \{1, 2, \ldots, n\}$,

$$\Pr\left[\bigwedge_{i \in I} \mathcal{E}_i\right] = \prod_{i \in I} \Pr[\mathcal{E}_i].$$

Definition:

Random variables $X_1, X_2, \ldots, X_n$ are **mutually independent** if for any subset $I \subset [n]$ and any values x_i, where $i \in I$,

$$\Pr\left[\bigwedge_{i \in I} (X_i = x_i)\right] = \prod_{i \in I} \Pr[X_i = x_i].$$

k-wise Independence

Definition:
Events $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ are k-**wise independent** if for any subset $I \subseteq \{1, 2, \ldots, n\}$, with $|I| \leq k$

$$\Pr\left[\bigwedge_{i \in I} \mathcal{E}_i\right] = \prod_{i \in I} \Pr[\mathcal{E}_i].$$

Definition:
Random variables $X_1, X_2, \ldots, X_n$ are k-**wise independent** if for any subset $I \subset [n]$ and any values x_i, where $i \in I$, with $|I| \leq k$

$$\Pr\left[\bigwedge_{i \in I}(X_i = x_i)\right] = \prod_{i \in I} \Pr[X_i = x_i].$$

pairwise: 2-wise

2-wise Independent Bits

uniform & independent bits: (random source)

$$X_1, X_2, \ldots, X_m \in \{0,1\}$$

Goal: 2-wise independent uniform bits:

$$Y_1, Y_2, \ldots, Y_n \in \{0,1\} \qquad n \gg m$$

a	b	$a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	0

nonempty subsets:

$$\emptyset \neq S_1, S_2, \ldots, S_{2^m-1} \subseteq \{1,2,\ldots,m\}$$

$$Y_j = \bigoplus_{i \in S_j} X_i$$

uniform & independent bits: $X_1, X_2, \ldots, X_m \in \{0,1\}$

nonempty subsets: $S_1, S_2, \ldots, S_{2^m-1} \subseteq \{1,2,\ldots,m\}$

$$Y_j = \bigoplus_{i \in S_j} X_i$$

2-wise independent uniform bits:

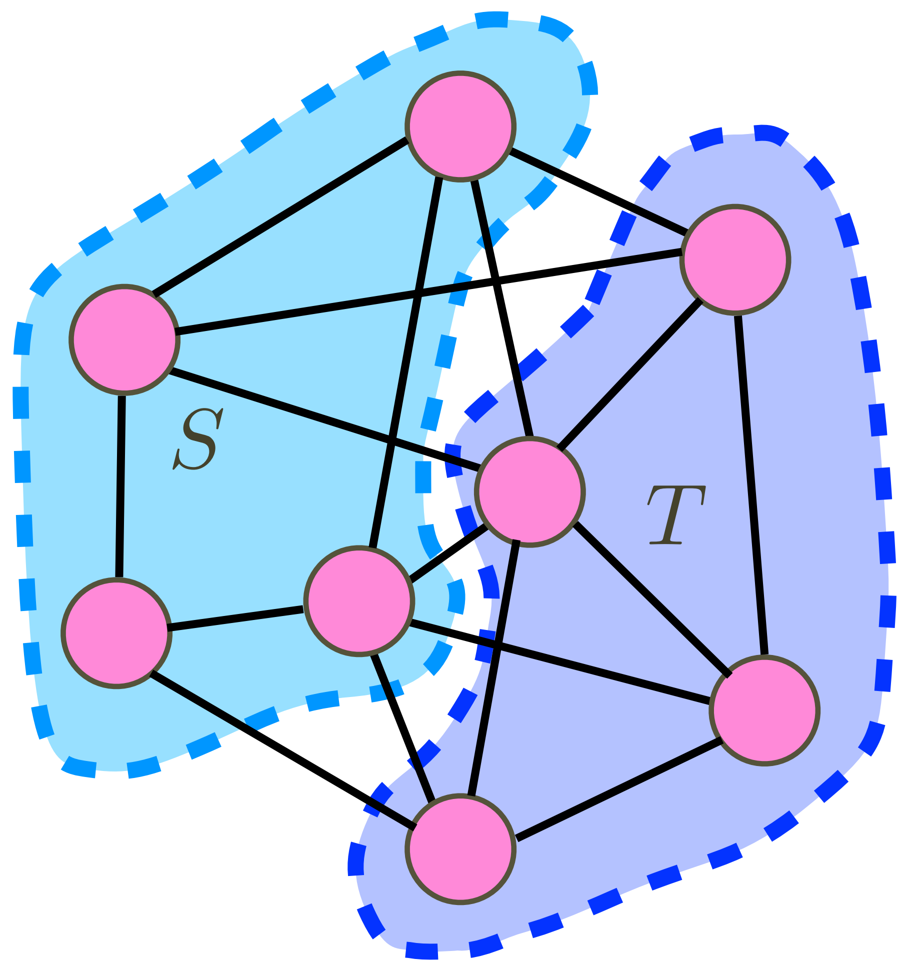
$$Y_1, Y_2, \ldots, Y_{2^m-1} \in \{0,1\}$$

$\log_2 n$ total random bits

⇨ n-1 pairwise independent bits

Max-Cut

- *partition* V into two parts: S and T
- maximize the cut $|C(S,T)|$
- NP-hard
  - 0.878~-approximation in poly-time by SDP
  - easy 0.5-approximation



$$C(S,T) = \{uv \in E \mid u \in S \text{ and } v \in T\}$$

Random Cut

for each vertex $v \in V$

uniform & independent $Y_v \in \{0,1\}$

$$Y_v = 1 \Rightarrow v \in S$$

$$Y_v = 0 \Rightarrow v \in T$$

for each edge $uv \in E$

$$Y_{uv} = \begin{cases} 1 & Y_u \neq Y_v \\ 0 & Y_u = Y_v \end{cases} \qquad |C(S,T)| = \sum_{uv \in E} Y_{uv}$$

$$\mathbf{E}[|C(S,T)|] = \sum_{uv \in E} \Pr[Y_u \neq Y_v] = \frac{|E|}{2} \geq \frac{OPT}{2}$$

Random Cut

for each vertex $v \in V$

uniform & 2-wise independent $Y_v \in \{0,1\}$

$Y_v = 1 \Rightarrow v \in S$

$Y_v = 0 \Rightarrow v \in T$

for each edge $uv \in E$

$$Y_{uv} = \begin{cases} 1 & Y_u \neq Y_v \\ 0 & Y_u = Y_v \end{cases} \qquad |C(S,T)| = \sum_{uv \in E} Y_{uv}$$

$$\mathbf{E}[|C(S,T)|] = \sum_{uv \in E} \Pr[Y_u \neq Y_v] = \frac{|E|}{2} \geq \frac{OPT}{2}$$

Derandomization

for each vertex $v \in V$

uniform & 2-wise independent $Y_v \in \{0,1\}$

$$Y_v = 1 \Rightarrow v \in S$$

$$Y_v = 0 \Rightarrow v \in T$$

for each edge $uv \in E$

$$\mathbf{E}[|C(S,T)|] = \sum_{uv \in E} \Pr[Y_u \neq Y_v] = \frac{|E|}{2} \geq \frac{OPT}{2}$$

$$V = \{v_1, v_2, \ldots, v_n\}$$

$Y_{v_1}, Y_{v_2}, \ldots, Y_{v_n}$ constructed from $\lceil \log_2(n+1) \rceil$ bits

try all $2^{\lceil \log_2(n+1) \rceil} = O(n^2)$ possibilities!

2-wise Independent Variables

random source: uniform and independent

$$X_0, X_1 \in [p]$$

Goal: uniform and 2-wise independent

$$Y_0, Y_1, \ldots, Y_{p-1} \in [p]$$ prime p

for $i \in [p]$ $\quad Y_i = (X_0 + i \cdot X_1) \bmod p$

uniformity: $\forall i, a \in [p] \quad \Pr[Y_i = a] = \frac{1}{p}$

2-wise independence: $\forall i \neq j, a, b \in [p]$

$$\Pr[Y_i = a \wedge Y_j = b] = \frac{1}{p^2}$$

uniform and independent $X_0, X_1 \in [p]$

for $i \in [p]$ $\quad Y_i = (X_0 + i \cdot X_1) \bmod p$

uniformity: $\quad \forall i, a \in [p]$

$$\Pr[Y_i = a]$$

$$= \Pr\left[(X_0 + i \cdot X_1) \bmod p = a\right]$$

$$= \sum_{j \in [p]} \Pr[X_1 = j] \cdot \Pr\left[(X_0 + ij) \bmod p = a\right]$$

$$= \frac{1}{p} \sum_{j \in [p]} \Pr\left[X_0 \equiv (a - ij) \pmod p\right]$$

$$= \frac{1}{p}$$

uniform and independent $X_0, X_1 \in [p]$

for $i \in [p]$ $\quad Y_i = (X_0 + i \cdot X_1) \bmod p$

2-wise independence: $\forall i \neq j, a, b \in [p]$

$$\Pr[Y_i = a \wedge Y_j = b]$$

$$= \Pr[(X_0 + iX_1) \bmod p = a \wedge (X_0 + jX_1) \bmod p = b]$$

$$\begin{cases} (X_0 + iX_1) \equiv a \pmod{p} \\ (X_0 + jX_1) \equiv b \pmod{p} \end{cases}$$

has unique solution $X_0 = x_0, X_1 = x_1$

$$= \Pr[X_0 = x_0 \wedge X_1 = x_1] \quad = \frac{1}{p^2}$$

Perfect Hashing

$S = \{ a, b, c, d, e, f \} \subseteq [N]$

uniform random h $[N] \to [m]$ Pr[*perfect*] > 1/2

e	b		d		f		c	a	

Table T:

$m = O(n^2)$

birthday!

UHA: Uniform Hash Assumption

search(x): retrieve h;

check whether $T[h(x)] = x$;

Universal Hash Family

(Carter-Wegman 1977)

universe $[N]$ range $[m]$

hash family $\mathcal{H}$ $\forall h \in \mathcal{H}$ $h : [N] \to [m]$

$\mathcal{H}$ is 2-universal if for uniform random $h \in \mathcal{H}$

$\forall$ distinct $x_1, x_2 \in [N]$

$$\Pr[h(x_1) = h(x_2)] \le \frac{1}{m}$$

"locally" like a uniform random hash function!

Universal Hash Family

(Carter-Wegman 1977)

universe $[N]$ range $[m]$

hash family $\mathcal{H}$ $\forall h \in \mathcal{H}$ $h : [N] \to [m]$

$\mathcal{H}$ is k-universal if for uniform random $h \in \mathcal{H}$

$\forall$ distinct $x_1, x_2, \ldots, x_k \in [N]$

$$\Pr[h(x_1) = h(x_2) = \cdots = h(x_k)] \leq \frac{1}{m^{k-1}}$$

"locally" like a uniform random hash function!

2-Universal $\mathcal{H}$

prime p for $a, b \in [p]$ define $h_{a,b} : [p] \to [p]$

$$h_{a,b}(x) = (a \cdot x + b) \bmod p$$

hash family $\mathcal{H} = \{h_{a,b} \mid a, b \in [p]\}$

$\mathcal{H}$ is 2-universal

$x_1 \neq x_2$ random $a, b \in [p]$

$h_{a,b}(x_1)$ and $h_{a,b}(x_2)$ are 2-wise independent

2-Universal $\mathcal{H}$

universe $[N]$ range $[m]$ prime $p \geq N$

for $a, b \in [p]$ define $h_{a,b} : [N] \to [m]$

$$h_{a,b}(x) = ((a \cdot x + b) \bmod p) \bmod m$$

hash family $\mathcal{H} = \{h_{a,b} \mid 1 \leq a \leq p-1, b \in [p]\}$

$\mathcal{H}$ is 2-universal

2-Universal $\mathcal{H}$

universe $[N]$ range $[m]$ prime $p \geq N$

for $a, b \in [p]$ define $h_{a,b} : [N] \to [m]$

$$h_{a,b}(x) = ((a \cdot x + b) \bmod p) \bmod m$$

hash family $\mathcal{H} = \{h_{a,b} \mid 1 \leq a \leq p-1, b \in [p]\}$

$\mathcal{H}$ is 2-universal

$x_1 \neq x_2$ random $1 \leq a \leq p-1, b \in [p]$

$$\Pr[h_{a,b}(x_1) = h_{a,b}(x_2)] = \frac{|\{(a,b) \mid h_{a,b}(x_1) = h_{a,b}(x_2)\}|}{p(p-1)}$$

for $a, b \in [p]$ **define** $h_{a,b} : [N] \to [m]$

$$h_{a,b}(x) = ((a \cdot x + b) \bmod p) \bmod m$$

$x_1 \neq x_2$ **random** $1 \leq a \leq p-1, b \in [p]$

$$\Pr[h_{a,b}(x_1) = h_{a,b}(x_2)] = \frac{|\{(a,b) \mid h_{a,b}(x_1) = h_{a,b}(x_2)\}|}{p(p-1)} \leq \frac{1}{m}$$

observation: $(a \cdot x_1 + b) \bmod p \neq (a \cdot x_2 + b) \bmod p$

$$\Rightarrow \begin{cases} (a \cdot x_1 + b) \bmod p = u \\ (a \cdot x_2 + b) \bmod p = v \end{cases} \qquad u \neq v$$

each (u,v) corresponds to exact one (a,b)

$$|\{(a,b) \mid h_{a,b}(x_1) = h_{a,b}(x_2)\}|$$

$$= |\{(u,v) \mid u \neq v, u \equiv v \pmod{m}\}| \leq p(p-1)/m$$

universe $[N]$ range $[m]$ prime $p \geq N$

for $a, b \in [p]$ define $h_{a,b} : [N] \to [m]$

$$h_{a,b}(x) = ((a \cdot x + b) \bmod p) \bmod m$$

hash family $\mathcal{H} = \{h_{a,b} \mid 1 \leq a \leq p-1, b \in [p]\}$

$\mathcal{H}$ is 2-universal

$\forall$ distinct $x_1, x_2, \ldots, x_n \in [N]$

uniform random $h \in \mathcal{H}$

$$\forall i \neq j \qquad \Pr[h(x_i) = h(x_j)] \leq \frac{1}{m}$$

Collision Number

$\mathcal{H}$ is 2-universal

uniform random $h \in \mathcal{H}$

$\forall$ distinct $x_1, x_2, \ldots, x_n \in [N]$

$$\forall i \neq j \qquad \Pr[h(x_i) = h(x_j)] \leq \frac{1}{m}$$

$$\forall i \neq j \qquad X_{ij} = \begin{cases} 1 & h(x_i) = h(x_j) \\ 0 & h(x_i) \neq h(x_j) \end{cases}$$ collision

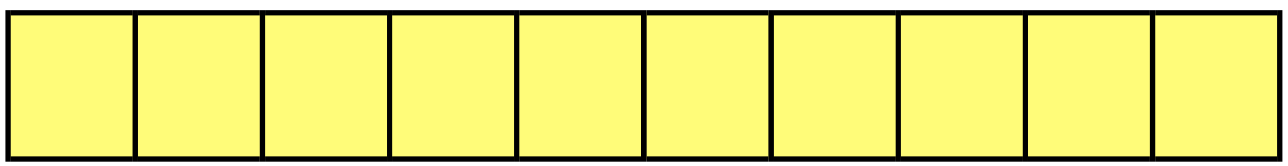
collision no.: $X = \sum_{i<j} X_{ij} \qquad \mathbf{E}[X] = \sum_{i<j} \mathbf{E}[X_{ij}] \leq \frac{n^2}{2m}$

birthday: $\Pr[X \geq 1] \leq \frac{1}{\mathbf{E}[X]} \leq \frac{2m}{n^2} = \frac{1}{2} \qquad n = 2\sqrt{m}$

Perfect Hashing

$$S = \{x_1, x_2, \ldots, x_n\} \subseteq [N]$$

2-universal h $[N] \to [m]$ Pr[*perfect*] > 1/2

Table T: 

$m = O(n^2)$

birthday!

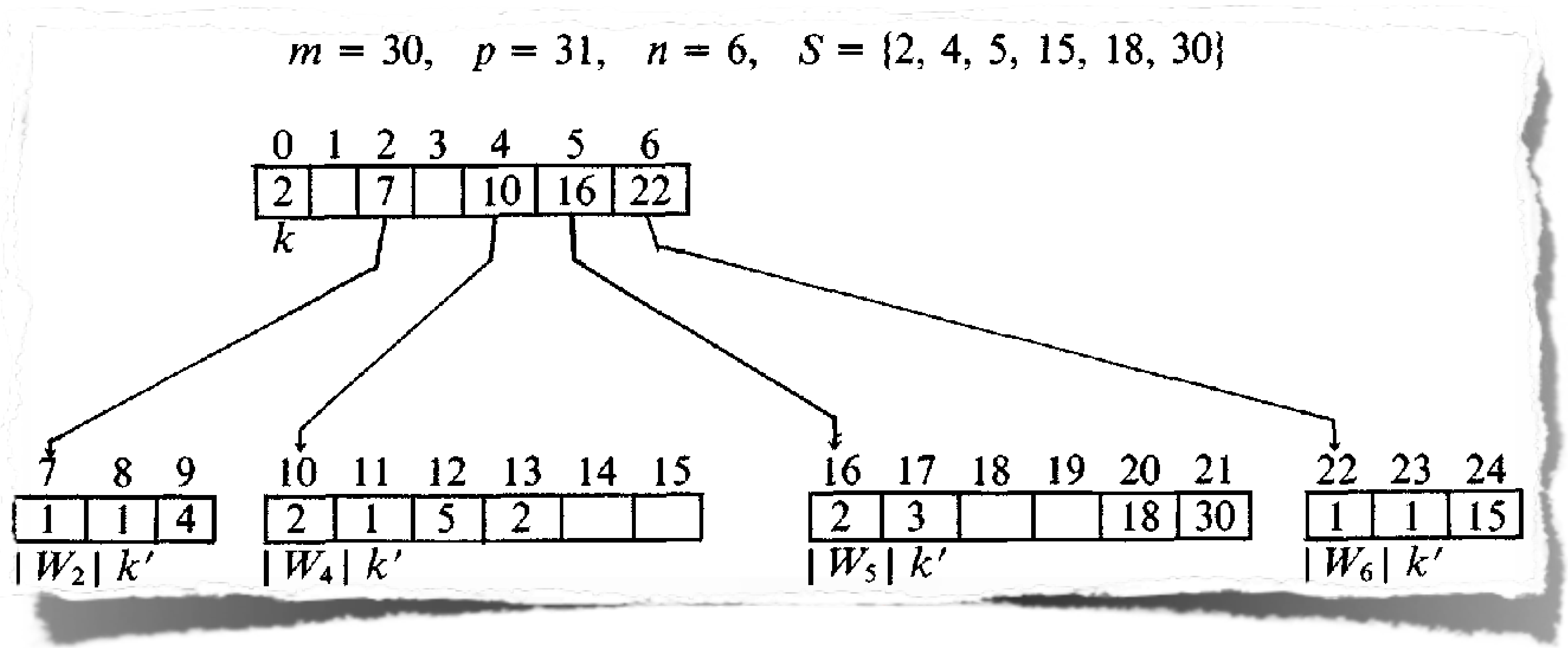
h is from 2-universal $\mathcal{H}$

search(x): retrieve h;
check whether $T[h(x)] = x$;

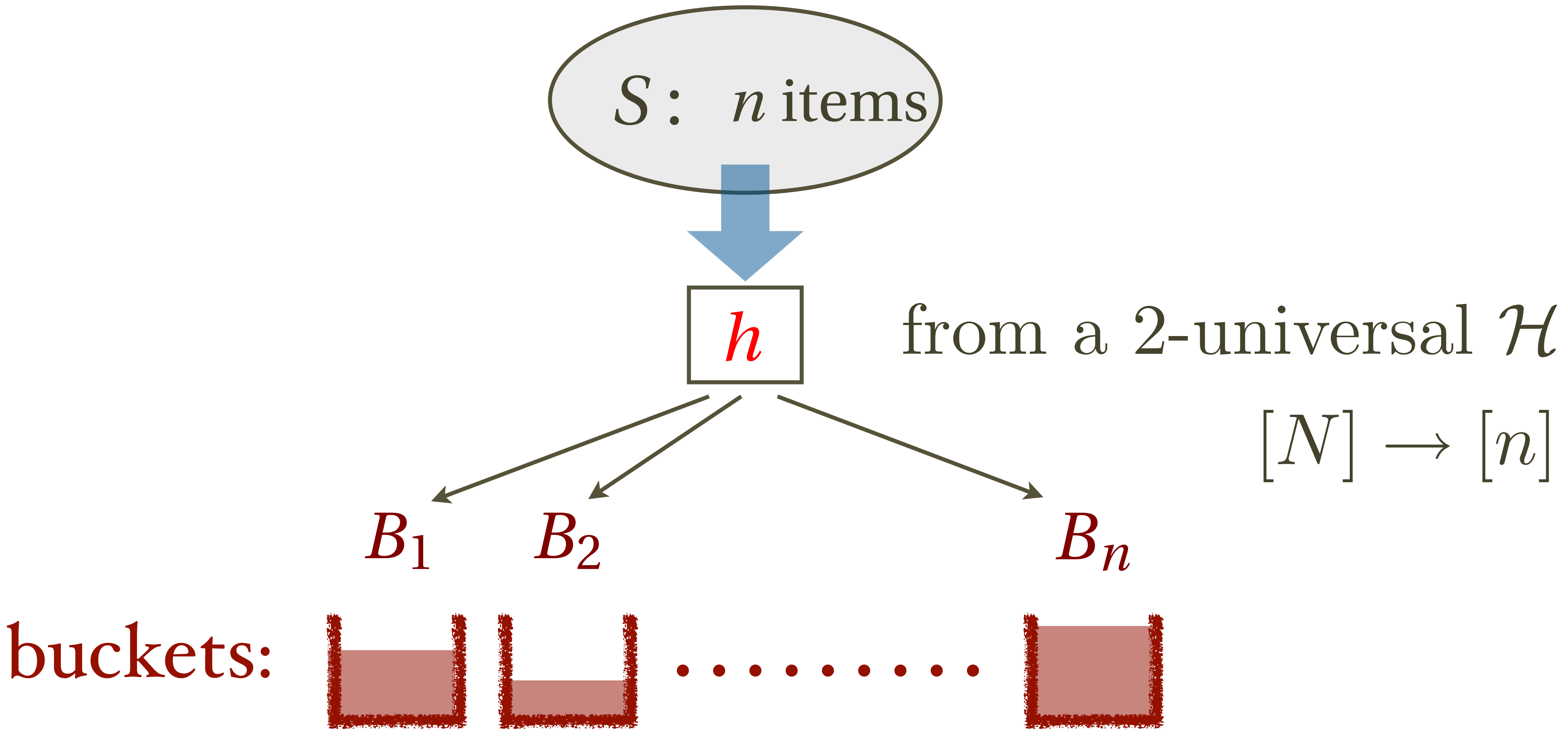
FKS Perfect Hashing

(Fredman, Komlós, Szemerédi, 1984)

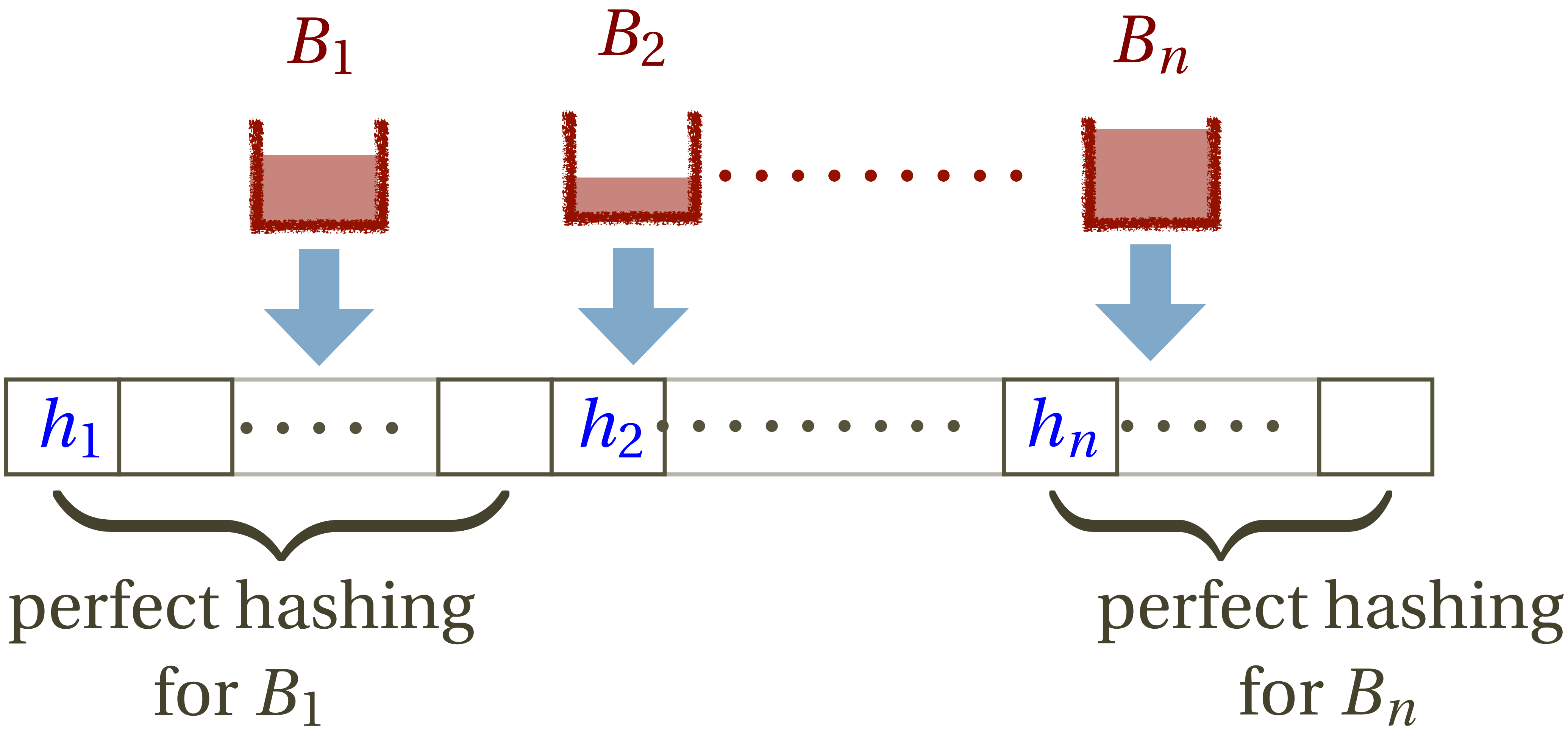
Goal: $O(n)$ space, $O(1)$ worst-case search time



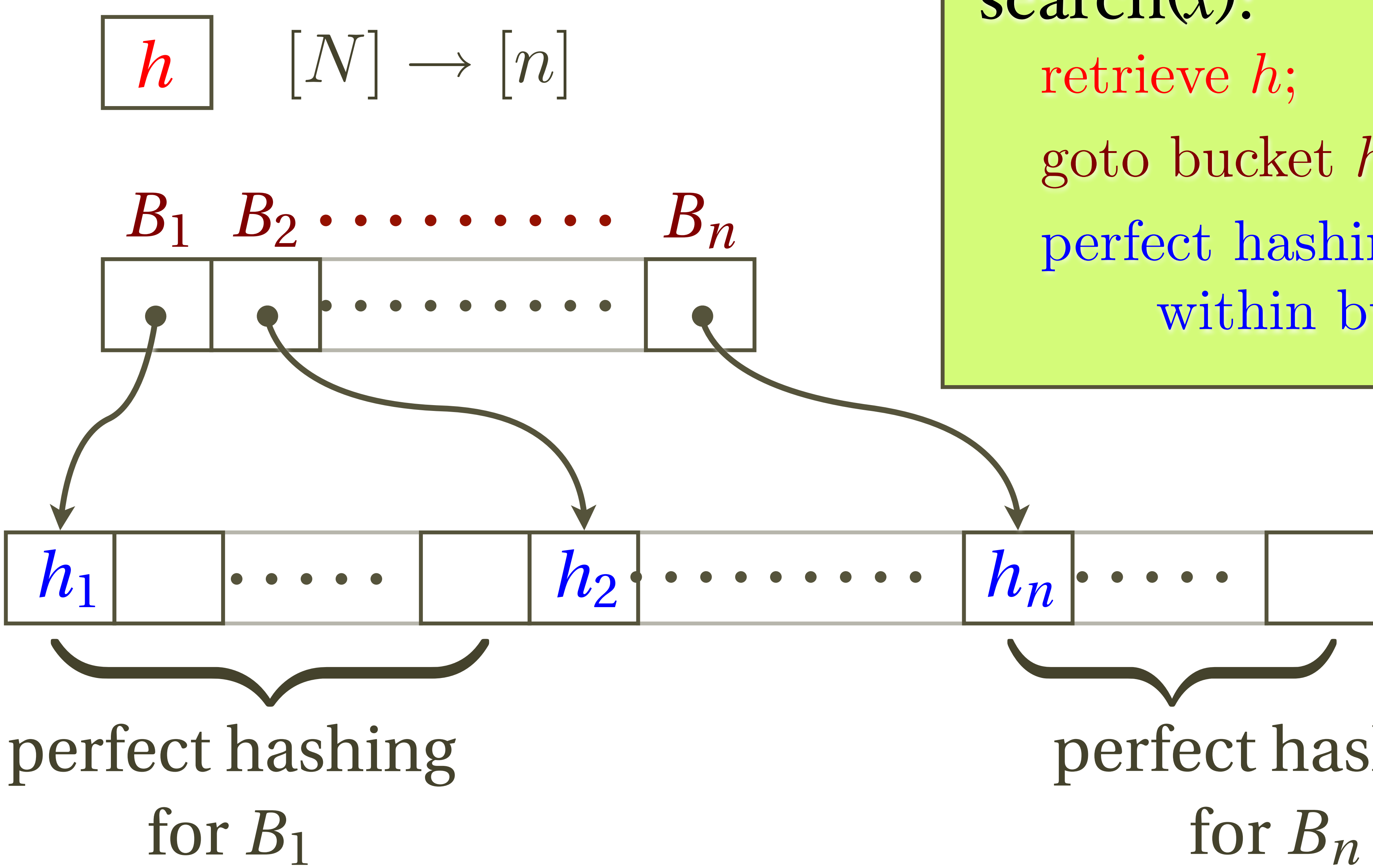
FKS Perfect Hashing



FKS Perfect Hashing



FKS Perfect Hashing



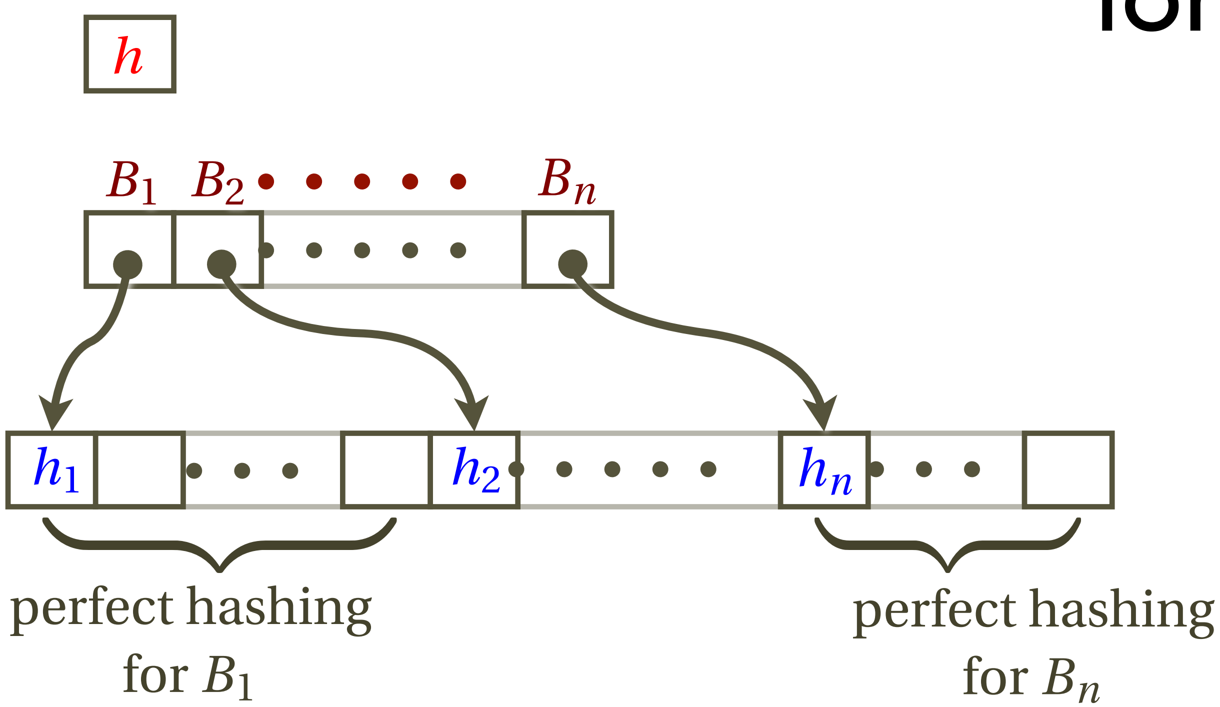
search(x):

retrieve h;

goto bucket $h(x)$;

perfect hashing within bucket;

FKS Perfect Hashing

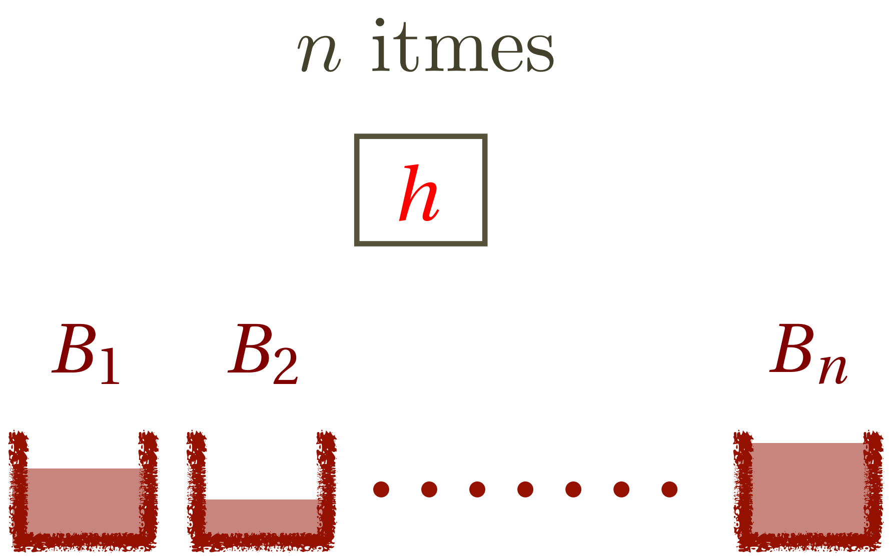


for a set S of n items:

- search time: $O(1)$
- space ?

Goal: $$\sum_{i=1}^{n} |B_i|^2 = O(n)$$

FKS Perfect Hashing



for a set S of n items:

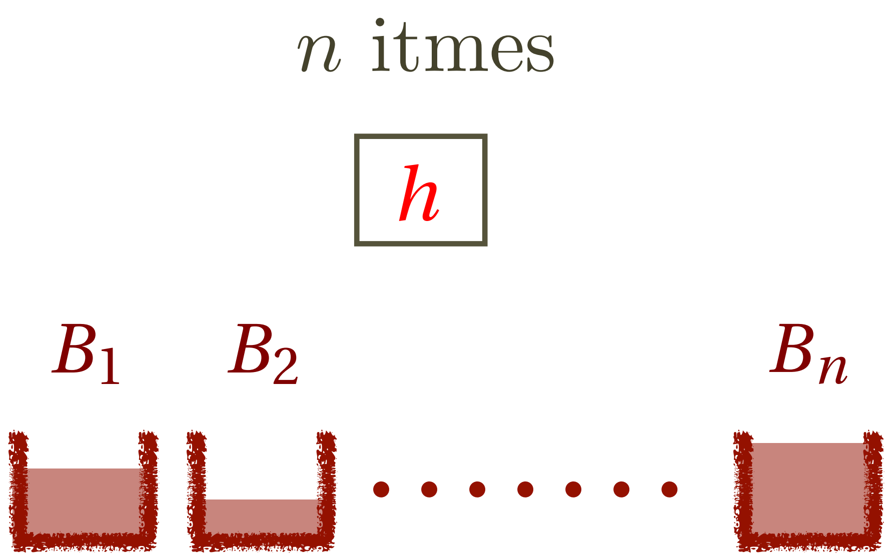
uniform random $h \in \mathcal{H}$

$$\sum_{i=1}^{n} |B_i|^2 = O(n)$$

Collision #: $\displaystyle\sum_{i=1}^{n} \binom{|B_i|}{2} = \frac{1}{2} \sum_{i=1}^{n} |B_i|(|B_i| - 1)$

$$= \frac{1}{2} \left(\sum_{i=1}^{n} |B_i|^2 - \sum_{i=1}^{n} |B_i| \right) = \frac{1}{2} \left(\sum_{i=1}^{n} |B_i|^2 - n \right)$$

FKS Perfect Hashing



for a set S of n items:

uniform random $h \in \mathcal{H}$

$$\sum_{i=1}^{n} |B_i|^2 = O(n)$$

$$\sum_{i=1}^{n} |B_i|^2 = n + 2 \cdot (\text{collision } \#)$$

$$\mathbf{E}[\text{collision } \#] \leq \frac{n}{2}$$

$$\mathbf{E}\left[\sum_{i=1}^{n} |B_i|^2\right] \leq 2n$$

Markov !

$\leq 4n$ with prob $\geq \frac{1}{2}$.

FKS Perfect Hashing

(Fredman, Komlós, Szemerédi, 1984)

Goal: $O(n)$ space, $O(1)$ worst-case search time