Hashing

KT Secs 13.6

Hashing universal

U: conjunto universo (contém todas as possíveis chaves).

n: um número muito menor que $|U|$.

$\mathcal{H}$: conjunto de funções de U em $\{0, \ldots, n-1\}$.

Hashing universal

U: conjunto universo (contém todas as possíveis chaves).

n: um número muito menor que $|U|$.

$\mathcal{H}$: conjunto de funções de U em $\{0, \ldots, n-1\}$.

$\mathcal{H}$ é uma coleção universal de hashing se,
para cada par de chaves k, ℓ em U,
o número de funções h em $\mathcal{H}$ tais que $h(k) = h(\ell)$
é no máximo $|\mathcal{H}|/n$.

Hashing universal

U: conjunto universo (contém todas as possíveis chaves).

n: um número muito menor que $|U|$.

$\mathcal{H}$: conjunto de funções de U em $\{0, \ldots, n-1\}$.

$\mathcal{H}$ é uma coleção universal de hashing se,
para cada par de chaves k, ℓ em U,
o número de funções h em $\mathcal{H}$ tais que $h(k) = h(\ell)$
é no máximo $|\mathcal{H}|/n$.

Fixe $k, \ell \in U$.

O que acontece se escolhermos uma h em $\mathcal{H}$
aleatoriamente com probabilidade uniforme?

Hashing universal

U: conjunto universo (contém todas as possíveis chaves).

n: um número muito menor que $|U|$.

$\mathcal{H}$: conjunto de funções de U em $\{0, \ldots, n-1\}$.

$\mathcal{H}$ é uma coleção universal de hashing se,
para cada par de chaves k, ℓ em U,
o número de funções h em $\mathcal{H}$ tais que $h(k) = h(\ell)$
é no máximo $|\mathcal{H}|/n$.

Fixe $k, \ell \in U$.

O que acontece se escolhermos uma h em $\mathcal{H}$
aleatoriamente com probabilidade uniforme?

Qual é a chance de $h(k) = h(\ell)$?

Hashing universal

n: um número muito menor que $|U|$.

$\mathcal{H}$: conjunto de funções de U em $\{0, \ldots, n-1\}$.

$\mathcal{H}$ é uma coleção universal de hashing se,
para cada par de chaves k, ℓ em U, o número de
funções h em $\mathcal{H}$ tais que $h(k) = h(\ell)$ é no máximo $|\mathcal{H}|/n$.

Fixe $k, \ell \in U$.

O que acontece se escolhermos uma h em $\mathcal{H}$
aleatoriamente com probabilidade uniforme?

Qual é a chance de $h(k) = h(\ell)$?

É, dentre todas as $|\mathcal{H}|$ funções h, escolhermos uma
das no máximo $|\mathcal{H}|/n$ para as quais vale a igualdade.

Hashing universal

n: um número muito menor que $|U|$.

$\mathcal{H}$: conjunto de funções de U em $\{0, \ldots, n-1\}$.

$\mathcal{H}$ é uma coleção universal de hashing se,
para cada par de chaves k, ℓ em U, o número de
funções h em $\mathcal{H}$ tais que $h(k) = h(\ell)$ é no máximo $|\mathcal{H}|/n$.

Fixe $k, \ell \in U$.

O que acontece se escolhermos uma h em $\mathcal{H}$
aleatoriamente com probabilidade uniforme?

Qual é a chance de $h(k) = h(\ell)$?

É, dentre todas as $|\mathcal{H}|$ funções h, escolhermos uma
das no máximo $|\mathcal{H}|/n$ para as quais vale a igualdade.
Ou seja, é no máximo $1/n$.

Formalizando...

Teorema: Seja $S \subseteq U$ tal que $|S| \leq |U|$ e $u \in U$.
Se h é escolhida aleatoriamente de uma coleção universal $\mathcal{H}$ e X é o número de elementos s em S tais que $h(s) = h(u)$, então $\mathrm{E}[X] \leq 1$ se $u \notin S$ e $\mathrm{E}[X] \leq 2$ se $u \in S$.

Formalizando...

Teorema: Seja $S \subseteq U$ tal que $|S| \leq |U|$ e $u \in U$.
Se h é escolhida aleatoriamente de uma coleção universal $\mathcal{H}$ e X é o número de elementos s em S tais que $h(s) = h(u)$, então $\mathrm{E}[X] \leq 1$ se $u \notin S$ e $\mathrm{E}[X] \leq 2$ se $u \in S$.

Prova: Lembre-se que $n = |U|$.

Seja X_s a variável binária que vale 1 se $h(s) = h(u)$.

Formalizando...

Teorema: Seja $S \subseteq U$ tal que $|S| \leq |U|$ e $u \in U$.
Se h é escolhida aleatoriamente de uma coleção universal $\mathcal{H}$ e X é o número de elementos s em S tais que $h(s) = h(u)$, então $\mathrm{E}[X] \leq 1$ se $u \notin S$ e $\mathrm{E}[X] \leq 2$ se $u \in S$.

Prova: Lembre-se que $n = |U|$.

Seja X_s a variável binária que vale 1 se $h(s) = h(u)$.

Note que $X = \sum_s X_s$.

Formalizando...

Teorema: Seja $S \subseteq U$ tal que $|S| \leq |U|$ e $u \in U$.
Se h é escolhida aleatoriamente de uma coleção universal $\mathcal{H}$ e X é o número de elementos s em S tais que $h(s) = h(u)$, então $\mathrm{E}[X] \leq 1$ se $u \notin S$ e $\mathrm{E}[X] \leq 2$ se $u \in S$.

Prova: Lembre-se que $n = |U|$.

Seja X_s a variável binária que vale 1 se $h(s) = h(u)$.

Note que $X = \sum_s X_s$.

Então $\Pr\{X_u = 1\} = 1$ e $\Pr\{X_s = 1\} \leq \frac{1}{n}$ se $s \neq u$.

Formalizando...

Teorema: Seja $S \subseteq U$ tal que $|S| \leq |U|$ e $u \in U$.
Se h é escolhida aleatoriamente de uma coleção universal $\mathcal{H}$ e X é o número de elementos s em S tais que $h(s) = h(u)$, então $\mathrm{E}[X] \leq 1$ se $u \notin S$ e $\mathrm{E}[X] \leq 2$ se $u \in S$.

Prova: Lembre-se que $n = |U|$.

Seja X_s a variável binária que vale 1 se $h(s) = h(u)$.

Note que $X = \sum_s X_s$.

Então $\Pr\{X_u = 1\} = 1$ e $\Pr\{X_s = 1\} \leq \dfrac{1}{n}$ se $s \neq u$.

Logo
$$\mathrm{E}[X] = \sum_{s \in S} \mathrm{E}[X_s] \leq \sum_{s \in S} \frac{1}{n} = \frac{|S|}{n} \leq 1 \quad \text{se } u \notin S$$

Formalizando...

Teorema: Seja $S \subseteq U$ tal que $|S| \leq |U|$ e $u \in U$.
Se h é escolhida aleatoriamente de uma coleção universal $\mathcal{H}$
e X é o número de elementos s em S tais que $h(s) = h(u)$,
então $\mathrm{E}[X] \leq 1$ se $u \notin S$ e $\mathrm{E}[X] \leq 2$ se $u \in S$.

Prova: Lembre-se que $n = |U|$.

Seja X_s a variável binária que vale 1 se $h(s) = h(u)$.

Note que $X = \sum_s X_s$.

Então $\Pr\{X_u = 1\} = 1$ e $\Pr\{X_s = 1\} \leq \dfrac{1}{n}$ se $s \neq u$.

Logo
$$\mathrm{E}[X] = \sum_{s \in S} \mathrm{E}[X_s] \leq \sum_{s \in S} \frac{1}{n} = \frac{|S|}{n} \leq 1 \quad \text{se } u \notin S$$

e
$$\mathrm{E}[X] = 1 + \sum_{s \in S \setminus \{u\}} \frac{1}{n} < 1 + \frac{|S|}{n} \leq 2 \quad \text{se } u \in S. \quad \square$$

Exemplo de coleção universal de hashing

Seja p um primo tal que $U \subseteq \{0, \ldots, p-1\}$.

$\mathbb{Z}_p$: conjunto $\{0, \ldots, p-1\}$.
$\mathbb{Z}_p^*$: conjunto $\{1, \ldots, p-1\}$.

Exemplo de coleção universal de hashing

Seja p um primo tal que $U \subseteq \{0, \ldots, p-1\}$.

$\mathbb{Z}_p$: conjunto $\{0, \ldots, p-1\}$.
$\mathbb{Z}_p^*$: conjunto $\{1, \ldots, p-1\}$.

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) = ((ak + b) \bmod p) \bmod n,$$

para todo k em U.

Exemplo de coleção universal de hashing

Seja p um primo tal que $U \subseteq \{0, \ldots, p-1\}$.

$\mathbb{Z}_p$: conjunto $\{0, \ldots, p-1\}$.
$\mathbb{Z}_p^*$: conjunto $\{1, \ldots, p-1\}$.

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n,$$

para todo k em U.

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Exemplo de coleção universal de hashing

Seja p um primo tal que $U \subseteq \{0, \ldots, p-1\}$.

$\mathbb{Z}_p$: conjunto $\{0, \ldots, p-1\}$.
$\mathbb{Z}_p^*$: conjunto $\{1, \ldots, p-1\}$.

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n,$$

para todo k em U.

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Note que $|\mathcal{H}| = p(p-1)$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam k, $\ell \in U$ tais que $k \neq \ell$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam k, $\ell \in U$ tais que $k \neq \ell$.

Queremos determinar quantas $h_{a,b} \in \mathcal{H}$ são tais que $h_{a,b}(k) = h_{a,b}(\ell)$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam k, $\ell \in U$ tais que $k \neq \ell$.

Queremos determinar quantas $h_{a,b} \in \mathcal{H}$ são tais que $h_{a,b}(k) = h_{a,b}(\ell)$.

Se $(ak + b) \bmod p = (a\ell + b) \bmod p$,
então $a(k - \ell) = 0 \bmod p$, o que implica que $a = 0$ ou $k = \ell$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam k, $\ell \in U$ tais que $k \neq \ell$.

Queremos determinar quantas $h_{a,b} \in \mathcal{H}$ são tais que $h_{a,b}(k) = h_{a,b}(\ell)$.

Se $(ak + b) \bmod p = (a\ell + b) \bmod p$,
então $a(k - \ell) = 0 \bmod p$, o que implica que $a = 0$ ou $k = \ell$.

Como $a \neq 0$ e $k \neq \ell$, $\quad r = (ak + b) \bmod p \neq (a\ell + b) \bmod p = s$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam $k, \ell \in U$ tais que $k \neq \ell$.

Queremos determinar quantas $h_{a,b} \in \mathcal{H}$ são tais que $h_{a,b}(k) = h_{a,b}(\ell)$.

Se $(ak + b) \bmod p = (a\ell + b) \bmod p$,
então $a(k - \ell) = 0 \bmod p$, o que implica que $a = 0$ ou $k = \ell$.

Como $a \neq 0$ e $k \neq \ell$, $\quad r = (ak + b) \bmod p \neq (a\ell + b) \bmod p = s$.

Ademais, cada par (a, b) está associado a um par distinto (r, s),
com $r \neq s$, já que $a = (r - s)(k - \ell)^{-1} \bmod p$ e $b = (r - ak) \bmod p$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam k, $\ell \in U$ tais que $k \neq \ell$.

Queremos determinar quantas $h_{a,b} \in \mathcal{H}$ são tais que $h_{a,b}(k) = h_{a,b}(\ell)$.

Como $a \neq 0$ e $k \neq \ell$, $\quad r = (ak + b) \bmod p \neq (a\ell + b) \bmod p = s$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam k, $\ell \in U$ tais que $k \neq \ell$.

Queremos determinar quantas $h_{a,b} \in \mathcal{H}$ são tais que $h_{a,b}(k) = h_{a,b}(\ell)$.

Como $a \neq 0$ e $k \neq \ell$, $\quad r = (ak + b) \bmod p \neq (a\ell + b) \bmod p = s$.

Para cada r em $\mathbb{Z}_p$, temos $p - 1$ valores possíveis para s em $\mathbb{Z}_p \setminus \{r\}$.
Destes, não mais que $p/n - 1 \leq (p-1)/n$ são tais que $s = r \bmod n$.

Exemplo de coleção universal de hashing

Para todo a em $\mathbb{Z}_p^*$ e b em $\mathbb{Z}_p$, seja

$$h_{a,b}(k) \;=\; ((ak + b) \bmod p) \bmod n, \quad \text{para todo } k \text{ em } U.$$

A coleção $\mathcal{H} \;=\; \{h_{a,b} : a \in \mathbb{Z}_p^* \text{ e } \; b \in \mathbb{Z}_p\}$ é universal.

Esboço da prova: Sejam k, $\ell \in U$ tais que $k \neq \ell$.

Queremos determinar quantas $h_{a,b} \in \mathcal{H}$ são tais que $h_{a,b}(k) = h_{a,b}(\ell)$.

Como $a \neq 0$ e $k \neq \ell$, $\quad r = (ak + b) \bmod p \neq (a\ell + b) \bmod p = s$.

Para cada r em $\mathbb{Z}_p$, temos $p - 1$ valores possíveis para s em $\mathbb{Z}_p \setminus \{r\}$.
Destes, não mais que $p/n - 1 \leq (p-1)/n$ são tais que $s = r \bmod n$.

Ou seja, $h_{a,b}(k) = h_{a,b}(\ell)$
para não mais que $p(p-1)/n = |\mathcal{H}|/n$ das funções $h_{a,b}$ de $\mathcal{H}$. □