

Kiegészítés a 17. tételhez.

(A régi 17-es tételből hiányzott)

Hash függvények egyszerű egyenletestől univerzálisig

A hash függvény formája:

```
int hash( char *s, int n )
{
    int sum = 0;
    while( n-- ) sum = sum + *s++;
    return sum % 256;
}
```

ez a függvény a 0..255 egy értékét adja vissza.

Vagy: *xor* függvény is jó:

```
sum = sum ^ *s++;
```

Egyszerű egyenletes hasítás

$P(k)$ legyen annak a valószínűsége hogy a k kulcs előfordul, a hasító táblánkban m hely van, egy egyenletes has függvény, $h(k)$ biztosítani fogja:

$$\sum_{k|h(k)=0} P(k) = \sum_{k|h(k)=1} P(k) = \dots \sum_{k|h(k)=m-1} P(k) = \frac{1}{m}$$

(szumma minden k -ra, melyre $h(k) = 0$).

A kulcsok száma minden helyre azonos.

Ha a kulcsok a $[0, r)$ -en véletlenszerűen szétszórótt egészek, akkor

$$h(k) = \left\lfloor \frac{mk}{r} \right\rfloor$$

egy egyenletes hash függvény. A kulcsokat egészek egy intervallumára képeztük le: $0 \leq k < r$. Most csökkentsük ezt az intervallumot $[0, m)$ -re, ahol m a hash tábla egy elfogadható mérete.

Hogyan?

- Osztás (mod) -

$$h(k) = k \mod m$$

m mi legyen: pl. 2^n

- Szorzó módszer: konstanssal szorzunk ($0 < A < 1$), utána kivesszük a tört részt belőle:

$$(kA - \lfloor kA \rfloor)$$

Utána ezt még felszorozzuk m -mel:

$$h(k) = \lfloor m * (kA - \lfloor kA \rfloor) \rfloor$$

Itt jó a 2 hatvány.

- Univerzális hashelés

Univerzális hashelés

A hash függvényt véletlenül, az aktuálisan tárolandó kulcsoktól függetlenül választjuk meg.

A hasító függvényt egy függvényosztályból futás közben véletlenül választjuk ki.

Legyen H a hasító függvények véges halmaza, ezek egy K kulcs univerzumot a $[0, m)$ tartományba képeznek le.

H **univerzális**, ha $\forall x, y \in K, x \neq y$ kulcspárra azoknak a $h \in H$ hasító függvényeknek a száma, melyre $h(x) = h(y)$, pontosan $\frac{|H|}{m}$. Ilyenkor a kulcsütközés valószínűsége pontosan $\frac{1}{m}$

Megválasztásuk. Vegyünk egy olyan p prím számot, mely elég nagy hogy minden kulcs benne legyen a $[0, \dots, p-1]$ -ben ($p > m$).

Jelölés:

$$Z_p = \{0, 1, \dots, p-1\}, Z_p^* = \{1, 2, \dots, p-1\}$$

Definiáljuk: $\forall a \in Z_p^*, \forall b \in Z_p$:

$$h_{a,b}(k) = ((a * k + b) \mod p) \mod m$$

Az ilyen függvények osztálya:

$$H_{p,m} = \{h_{a,b} : a \in Z_p^*, b \in Z_p\}$$

Tétel. A hasító függvények fenti egyenlőségekkel definiált $H_{p,m}$ osztálya univerzális.