DZ 1.

1. Ilustracija korištenja f-je raspršivanja

1. Demonstrirajte ubacivanje ključeva 77, 69, 39, 70, 6, 8, 40, 89, 49, 15 u hash tablicu veličine $m = 19$.

a) u kojem se kolizije rj. ulančavanjem, gdje je dana f-ja raspršenja $h(k) = k \bmod m$.

0	
1	77 → 39
2	40
3	
4	
5	
6	6
7	
8	8
9	
10	
11	49
12	69
13	70 → 89
14	
15	15
16	
17	
18	

$h(77) = 77 \bmod 19 = 1$
$h(69) = 12$
$h(39) = 1$
$h(70) = 13$
$h(6) = 6$
$h(8) = 8$
$h(40) = 2$
$h(89) = 13$
$h(49) = 11$
$h(15) = 15$

b) u kojem se kolizija rj. probiranjem za $i = 0, \dots, m-1$, koristeći dvostruko probiranje.

$h(k, i) = (h_1(k) + i \cdot h_2(k)) \bmod m$
$h_1(k) = k \bmod m$
$h_2(k) = 1 + (k \bmod 18)$

$h(77, 0) = (1 + 0 \cdot (1+5)) \bmod 19 = 1$
$h(69, 0) = (12 \bmod 19) = 12$
$h(39, 0) = 1$ X
$h(39, 1) = (1 + 1(1+3)) \bmod 19 = 5$
$h(70, 0) = 13 \bmod 19 = 13$
$h(6, 0) = 6$
$h(8, 0) = 8$
$h(40, 0) = 2$
$h(89, 0) = 13$ X
$h(89, 1) = (13 + 1 \cdot (1+17)) \bmod 19 = 12$ X
$h(89, 2) = (13 + 2 \cdot 18) \bmod 19 = 11$
$h(49, 0) = 11$ X
$h(49, 1) = (11 + 1 \cdot 14) \bmod 19 = 6$ X
$h(49, 2) = (11 + 2 \cdot 14) \bmod 19 = 1$ X
$h(49, 3) = (11 + 3 \cdot 14) \bmod 19 = 15$
$h(15, 0) = 15$
$h(15, 1) = (15 + 1 \cdot 16) \bmod 19 = 12$
$h(15, 2) = (15 + 2 \cdot 16) \bmod 19 = 9$

	0
77	1
40	2
	3
	4
39	5
6	6
	7
8	8
15	9
	10
89	11
69	12
70	13
	14
49	15
	16
	17
	18

2. Promatrite m-znamenkasti dec. br. $x_1 x_2 \dots x_m$ $(x_i \in \{0, \dots, 9\})$. Je li hash f-ja $f(x) = \sum_{i=1}^{m} a_i x_i \pmod 8$ univerzalna?

Za a_i, $i = 1, \dots, m$ nezavisne sluč. varijable uniformno izabrane iz $\{0, 1, 2, \dots, 7\}$? Obrazložite, ako je dajete razlog zašto je, ako nije dajte kontraprimjer.

Rj: $$f(x) = \sum_{i=1}^{m} a_i x_i \pmod 8$$

$$a_i \in \{0, 1, 2, 3, 4, 5, 6, 7\}$$

$$x_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

Pitamo se je li $\mathcal{H} = \{f\}$ uniformno heširanje

- Neka su $x = \langle x_0, \dots, x_m \rangle$ i $y = \langle y_0, \dots, y_m \rangle$ različiti ključevi ... razlikuju se u barem jednoj znamenci, BSO neka je to prva pozicija npr. x na prvoj poz. ima znamenku 2, a y 4

$$\left.\begin{array}{l} x = \_\ \ 2\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_ \\ \quad\ \ = \qquad\ \ =\ \ =\ \ =\ \ =\ \ =\ \ =\ \ = \\ y = \_\ \ 4\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_ \end{array}\right\} x \neq y$$

Ako je $f \in \mathcal{H} \Rightarrow f(x) = f(y)$?

$$\sum_{i=1}^{m} a_i x_i \pmod 8 = \sum_{i=1}^{m} a_i y_i \pmod 8$$

$$\sum_{i=1}^{m} a_i (x_i - y_i) = 0 \pmod 8$$

$$a_1 (x_1 - y_1) + \sum_{i=2}^{m} a_i (x_i - y_i) = 0 \pmod 8$$

$$a_1 (x_1 - y_1) = -\sum_{i=2}^{m} a_i (x_i - y_i) \pmod 8$$

- kako je $x_1 \neq y_1$ mora postojati inverz

$$a_1 = \left(-\sum_{i=2}^{m} a_i (x_i - y_i)\right)(x_1 - y_1)^{-1} \pmod 8$$

dakle, za bilo koji odabir $a_0, a_2, a_3, \dots$ točno jedan odabir a_1 čini koliziju x i y.

$\Rightarrow f$ univerzalan skup.

Zadatak 2. Pretp. da koristimo hash f-ju h da se raspriši n raz. ključeva u tabl. T duljine m. Uz pretp. uniformnog raspršivanja, koliki je oček. br. kolizije? Preciznije, koliki je oč. kardinalitet $\{\{k, l\} : h(k) = h(l)\}$

Rj:

Neka je X_i indikator sluč. var. $1 \leq i \leq n$

$$X_i = \begin{cases} 1, & h(i) = h(j), \text{ za } i \neq j \\ 0, & \text{inače} \end{cases}$$

kolizija ... ključ raspršeni na isto mjesto u T.

$$X = \sum_{i \neq j} X_{ij} \Rightarrow \text{broj kolizije}$$

$$E[X] = E\left[\sum_{i \neq j} X_{ij}\right] = E X = \sum_{i \neq j} \underbrace{E[X_{ij}]}_{= P\{h(i) = h(j) ;\ i \neq j\} = \frac{1}{m} = \frac{n-i}{m}}$$

$$E[X] = \sum_{i \neq j} \frac{1}{m} = \sum \frac{n-i}{m} = \frac{n^2 - n(n+1)/2}{m} = \frac{n^2 - n}{2m} \quad \left(= \frac{n(n-1)}{2m} \right).$$

Zadatak 3. Tablica raspršenja (hash tablicu) veličine m koristi se za spremanje n ključeva, gdje $n \leq m/2$. Neka je korišteno otvoreno adresiranje s probiranjem za rezolviranje kolizije.

1. Uz pretp. unif. raspr., pok. da za $i = 1, \dots, n$ vjerojatnost da i-ta ubacivanje zahtjeva strogo više od k probiranja najviše 2^{-k}.

Rj: Dokazali smo da je očekivani broj probiranja u neuspješnoj pretrazi $\frac{1}{1-\alpha}$. U tom dokazu pok. smo koliko je

$$P\{X > i\} \leq \left(\frac{n}{m}\right)^{i-1} = \alpha^{i-1} \quad \text{za } n < m$$

$\rightarrow$ imamo $n \leq m/2$

$$P\{X > k\} = P\{X \geq k+1\} \leq \left(\frac{n}{m}\right)^k \leq \left(\frac{m}{2m}\right)^k = 2^{-k}$$

2. Pok. da za $i = 1 \dots m$ vjerojatnost da i-to ubacivanje zahtjeva više nego $2\lg m$ probiranja jest $O(1/m^2)$

Rj: $$P\{X_{2\lg m} \geq i\} = \frac{m}{m} \cdot \frac{m-1}{m-1} \cdot \ldots \cdot \frac{m-(2\lg m-1)}{m-(2\lg m-1)} \leq \left(\frac{m}{m}\right)^{2\lg m}$$

$$\leq \left(\frac{1}{2}\right)^{2\lg m} = 2^{(-2\lg m)} = \frac{1}{m^2}$$

$$P\{X_{2\lg m}\} \leq \frac{1}{m^2} \Rightarrow O\left(\frac{1}{m^2}\right)$$

Neka sluč. var. X_i označava broj probiranja potrebnih za i-to ubacivanje i neka je sluč. var. $X = \max_{1 \leq i \leq m} X_i$ najveći broj probiranja potrebnih za bilo koje od m ubacivanja.

3. Pokažite da je $Pr\{X > 2\lg m\} = O(1/m)$

Rj: $$P\{X_1 \geq 2\lg m\} + P\{X_2 \geq 2\lg m\} + \ldots + P\{X_m \geq 2\lg m\}$$

$$= \sum_{i=1}^{m} P\{X_i \geq 2\lg m\} \leq \sum_{i=1}^{m} \frac{1}{m^2} = \cancel{m} \cdot \frac{1}{m^{\cancel{2}}} = \frac{1}{m}$$

$$P\{x > 2\lg m\} = O(1/m)$$

4. Pok. da je očekivana duljina $E[x]$ najdužeg niza probiranja $O(\lg m)$.

- najduža moguća duljina niza probiranja je m

$$E[x] \leq P\{X \leq 2\lg m\}\, 2\lg m + P\{X > \lg m\}\, m = \frac{m-1}{m}\, 2\lg m + \frac{1}{\cancel{m}} \cdot \cancel{m}$$

$$= \left(1 - \frac{1}{m}\right) 2\lg m + 1 = 2\lg m + 1 - 2\frac{\lg m}{m} \Rightarrow O(\lg m)$$