Adatstruktúrák és Algoritmusok

7. gyakorlat

Hasítótáblák, feloszt algoritmus

Hasítótábla

A hasítótábla azt a szituációt kezeli, amikor nagy a kulcstartomány, de csak kevés kulcsot kell kezelnünk.
Ekkor szükségünk van egy hasítófüggvényre. A **hasítófüggvény** kulcshoz memóriacímet rendel.
Mivel nagy a kulcstartomány (azaz sok lehetséges kulcs van), de csak kevés a tárterület, így szükségszerűen bekövetkezik az a probléma, hogy a hasítófüggvény több kulcshoz ugyanazt a memóriacímet rendeli. Ezt a jelenséget nevezzük **ütközés**nek.

Az ütközésfeloldás típusai

- **Közvetlen címzés**: ütközés feloldása: láncolt lista alkalmazásával történik. Tehát a táblázat rései memóriacímeket tartalmaznak. Ezek láncolt listák fejei. A dolgozó adataihoz a megfelelő láncolt listán keresztül lehet eljutni. Kulcsot törölni a láncolt listáknál tanult módon lehet.
- **Nyíltcímzéses hasítás** Tömbös implementációt használunk. Háromféle módszert tanulunk, ezek a következők:
    1. **lineáris kipróbálás**
    2. **négyzetes kipróbálás**
    3. **dupla hasítás**

Műveletek: beszúrás, keresés, törlés.
Állapotok: Sz (szabad), F (foglalt), T (törölt).

Beszúrás

Bármelyik módszert alkalmazzuk a fenti három közül (lineáris kipróbálás, négyzetes kipróbálás, dupla hasítás) van két függvényünk:

- a $h_0(k)$ amely a k kulcshoz memóriacímet rendel,
- illetve a $h(k,i)$, amelynek $i = 0$ esetén az értéke azonos a $h_0(k)$ értékével, míg $i = 1, 2, \ldots, m-1$ esetén a $h(k,i)$ függvény segítségével végig lehet vele ugrálni a k kulcs által meghatározott sorrendben a táblázat összes résén.

Kezdetben minden rés állapota Sz. Ekkor a $h_0(k)$ és a $h(k, 0)$ ad egy címet, ahová a k kulcsot beszúrjuk és a rés állapotát Sz-ről F-re állítjuk. Általában a beszúrás úgy történik, hogy a $h(k, i)$ függvény által előírt sorrendben addig ugrálunk résről résre, amíg Sz, vagy T állapotú résre találunk, ide beszúrjuk a kulcsot, majd a rés állapotát F-re állítjuk.

Keresés

A k kulcsú rekordot keressük. A $h_0(k)$ és a $h(k, i)$ függvények által meghatározott sorrendben F, vagy T állapotú réseken ugrálunk. Az F állapotú rések esetén ellenőrizni kell, hogy megtaláltuk-e az k kulcsú rekordot. T állapot esetén tovább kell folytatni a keresést. A keresés kétféle eredménnyel zárulhat:

- Valamely F állapotú rés esetén megtaláljuk a k kulcsú rekordot.
- Sz állapotú rést kapunk, ami azt jelenti, hogy a táblázatban nem szerepel az adott kulcsú rekord.
  A keresés művelete rávilágít arra, hogy miért van szükségünk az T állapotra.
  Ha a törlés alkalmával a rés állapotát F-ről Sz-re állítanánk, akkor ez az Sz állapotú rés olyan kulcs esetén is leállítaná a keresést "nincs a táblázatban" eredménnyel, amely szerepel a táblázatban.

Törlés

Megkeressük az adott kulcsú rekordot és az állapotát F-ről T-re állítjuk. Ez egy logikai törlés, az adatok fizikailag még jelen vannak, de az adott kulcsú rekord már nem elérhető. Fizikai törlésre akkor kerül sor, amikor rekordot szúrunk be erre a résre, ekkor a rés fizikailag felülíródik az új rekorddal, és a rés állapota T-ről F-re állítódik.

1. Lineáris kipróbálás

$$\boxed{\begin{aligned} h_0(k) &= k \mod m, \\ h(k,i) &= (h_0(k)+i) \mod m \qquad (i = 0, 1, \ldots, m-1). \end{aligned}}$$

A következő példákban $m = 7$ (a táblázat sorainak a száma).

Feladat

$k = 17, 127, 20, 45, 17T, 81, 127T, 3, 10$ (T=törlés). Végezzük el a kijelölt műveleteket lineáris kipróbálás esetén.
Az üres táblázat:

0	Sz	
1	Sz	
2	Sz	
3	Sz	
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 17$-et.
$h_0(17) = 17 \mod 7 = 3$
$h(17, 0) = (3 + 0) \mod 7 = 3$

0	Sz	
1	Sz	
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 127$-et.
$h_0(127) = 127 \mod 7 = 1$
$h(127, 0) = (1 + 0) \mod 7 = 1$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 20$-at.

$h_0(20) = 20 \mod 7 = 6$

$h(20, 0) = (6 + 0) \mod 7 = 6$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 45$-öt.

$h_0(45) = 45 \mod 7 = 3$

$h(45,0) = (3+0) \mod 7 = 3$

$h(45,1) = (3+1) \mod 7 = 4$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	~~Sz~~F	45
5	Sz	
6	~~Sz~~F	20

Töröljük a $k = 17$-et.

$\boxed{\text{17T}}$ $h_0(17) = 17 \mod 7 = 3$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	45
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 81$-et.

$h_0(81) = 81 \mod 7 = 4$

$h(81, 0) = (4 + 0) \mod 7 = 4$

$h(81, 1) = (4 + 1) \mod 7 = 5$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	45
5	~~Sz~~F	81
6	~~Sz~~F	20

Töröljük a $k = 127$-et.

$\boxed{127\text{T}}$ $h_0(127) = 127 \mod 7 = 1$

0	Sz	
1	~~SzF~~T	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	45
5	~~Sz~~F	81
6	~~Sz~~F	20

Beszúrjuk a $k = 3$-at.

$h_0(3) = 3 \mod 7 = 3$

$$h(3,0) = (3+0) \mod 7 = 3$$

0	Sz	
1	~~SzF~~T	127
2	Sz	
3	~~SzFT~~F	~~17~~3
4	~~Sz~~F	45
5	~~Sz~~F	81
6	~~Sz~~F	20

Beszúrjuk a $k = 10$-et.

$h_0(10) = 10 \mod 7 = 3$

$$h(10,0) = (3+0) \mod 7 = 3$$
$$h(10,1) = (3+1) \mod 7 = 4$$
$$h(10,2) = (3+2) \mod 7 = 5$$
$$h(10,3) = (3+3) \mod 7 = 6$$
$$h(10,4) = (3+4) \mod 7 = 0$$

0	~~Sz~~ F	10
1	~~Sz~~ ~~F~~ T	127
2	Sz	
3	~~Sz~~ ~~F~~ ~~T~~ F	~~17~~ 3
4	~~Sz~~ F	45
5	~~Sz~~ F	81
6	~~Sz~~ F	20

2. Négyzetes kipróbálás

$$\boxed{\begin{aligned} h_0(k) &= k \mod m, \\ h(k,i) &= \left(h_0(k) + \frac{i(i+1)}{2}\right) \mod m \qquad (i = 0, 1, \ldots, m-1). \end{aligned}}$$

Állapotok: Sz (szabad), F (foglalt), T (törölt).
$m = 7$ (a táblázat sorainak a száma).

Feladat

$k = 17, 127, 20, 45, 17T, 81, 127T, 3, 10$. Végezzük el a kijelölt műveleteket négyzetes kipróbálás esetén.

0	Sz	
1	Sz	
2	Sz	
3	Sz	
4	Sz	
5	Sz	
6	Sz	

0	Sz	
1	Sz	
2	Sz	
3	Sz	
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 17$-et.
$h_0(17) = 17 \mod 7 = 3$
$h(17, 0) = \left(3 + \frac{0 \cdot 1}{2}\right) \mod 7 = 3$

0	Sz	
1	Sz	
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 127$-et.
$h_0(127) = 127 \mod 7 = 1$
$h(127, 0) = \left(1 + \frac{0 \cdot 1}{2}\right) \mod 7 = 1$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 20$-at.
$h_0(20) = 20 \mod 7 = 6$
$h(20, 0) = \left(6 + \frac{0 \cdot 1}{2}\right) \mod 7 = 6$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 45$-öt.
$h_0(45) = 45 \mod 7 = 3$
$h(45, 0) = \left(3 + \frac{0 \cdot 1}{2}\right) \mod 7 = 3$
$h(45, 1) = \left(3 + \frac{1 \cdot 2}{2}\right) \mod 7 = 4$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	~~Sz~~F	45
5	Sz	
6	~~Sz~~F	20

Töröljük a $k = 17$-et.
$\boxed{17T}$ $h_0(17) = 17 \mod 7 = 3$
$h(17,0) = \left(3 + \frac{0 \cdot 1}{2}\right) \mod 7 = 3$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	45
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 81$-et.

$h_0(81) = 81 \mod 7 = 4$

$h(81, 0) = \left(4 + \frac{0 \cdot 1}{2}\right) \mod 7 = 4$

$h(81, 1) = \left(4 + \frac{1 \cdot 2}{2}\right) \mod 7 = 5$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	45
5	~~Sz~~F	81
6	~~Sz~~F	20

Töröljük a $k = 127$-et.

$\boxed{127\text{T}}$ $h_0(127) = 127 \mod 7 = 1$

$h(127, 0) = \left(1 + \frac{0 \cdot 1}{2}\right) \mod 7 = 1$

0	Sz	
1	~~SzF~~T	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	45
5	~~Sz~~F	81
6	~~Sz~~F	20

Beszúrjuk a $k = 3$-at.

$h_0(3) = 3 \mod 7 = 3$

$$h(3,0) = \left(3 + \frac{0 \cdot 1}{2}\right) \mod 7 = 3$$

0	Sz	
1	~~Sz~~ ~~F~~ T	127
2	Sz	
3	~~Sz~~ ~~F~~ ~~T~~ F	~~17~~ 3
4	~~Sz~~ F	45
5	~~Sz~~ F	81
6	~~Sz~~ F	20

Beszúrjuk a $k = 10$-et.

$h_0(10) = 10 \mod 7 = 3$

$$h(10,0) = \left(3 + \frac{0 \cdot 1}{2}\right) \mod 7 = 3$$
$$h(10,1) = \left(3 + \frac{1 \cdot 2}{2}\right) \mod 7 = 4$$
$$h(10,2) = \left(3 + \frac{2 \cdot 3}{2}\right) \mod 7 = 6$$
$$h(10,3) = \left(3 + \frac{3 \cdot 4}{2}\right) \mod 7 = 2$$

0	Sz	
1	~~Sz~~ ~~F~~ T	127
2	~~Sz~~ F	10
3	~~Sz~~ ~~F~~ ~~T~~ F	~~17~~ 3
4	~~Sz~~ F	45
5	~~Sz~~ F	81
6	~~Sz~~ F	20

3. Dupla hasítás

$$\boxed{\begin{aligned} h_0(k) &= k \mod m, \\ h_1(k) &= 1 + (k \mod (m-1)) \\ h(k,i) &= (h_0(k) + ih_1(k)) \mod m \qquad\qquad (i = 0, 1, \ldots, m-1). \end{aligned}}$$

$m = 7$ (a táblázat sorainak a száma).

Feladat

$k = 17, 127, 20, 45, 17T, 81, 127T, 3, 10$. Végezzük el a kijelölt műveleteket dupla hasítás segítségével.

0	Sz	
1	Sz	
2	Sz	
3	Sz	
4	Sz	
5	Sz	
6	Sz	

0	Sz	
1	Sz	
2	Sz	
3	Sz	
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 17$-et.

$$\left.\begin{array}{l} h_0(17) = 17 \mod 7 = 3 \\ h_1(17) = 1 + (17 \mod 6) = 6 \end{array}\right\} \quad \Rightarrow$$

$$h(17, 0) = (3 + 0 \cdot 6) \mod 7 = 3$$

0	Sz	
1	Sz	
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 127$-et.

$$\left.\begin{aligned} h_0(127) &= 127 \mod 7 = 1 \\ h_1(127) &= 1 + (127 \mod 6) = 2 \end{aligned}\right\} \Rightarrow$$

$$h(127, 0) = (1 + 0 \cdot 2) \mod 7 = 1$$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	Sz	

Beszúrjuk a $k = 20$-at.

$$\left.\begin{aligned} h_0(20) &= 20 \mod 7 = 6 \\ h_1(20) &= 1 + (20 \mod 6) = 3 \end{aligned}\right\} \quad \Rightarrow$$

$$h(20, 0) = (6 + 0 \cdot 3) \mod 7 = 6$$

0	Sz	
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 45$-öt.

$$\left.\begin{aligned} h_0(45) &= 45 \mod 7 = 3 \\ h_1(45) &= 1 + (45 \mod 6) = 4 \end{aligned}\right\} \quad \Rightarrow$$

$$h(45, 0) = (3 + 0 \cdot 4) \mod 7 = 3$$
$$h(45, 1) = (3 + 1 \cdot 4) \mod 7 = 0$$

0	~~Sz~~F	45
1	~~Sz~~F	127
2	Sz	
3	~~Sz~~F	17
4	Sz	
5	Sz	
6	~~Sz~~F	20

Töröljük a $k = 17$-et.

$$17T \quad \left.\begin{array}{l} h_0(17) = 17 \mod 7 = 3 \\ h_1(17) = 1 + (17 \mod 6) = 6 \end{array}\right\} \quad \Rightarrow$$

$$h(17, 0) = (3 + 0 \cdot 6) \mod 7 = 3$$

0	~~Sz~~F	45
1	~~Sz~~F	127
2	Sz	
3	~~SzF~~T	17
4	Sz	
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 81$-et.

$$\left.\begin{aligned} h_0(81) &= 81 \mod 7 = 4 \\ h_1(81) &= 1 + (81 \mod 6) = 4 \end{aligned}\right\} \quad \Rightarrow$$

$$h(81, 0) = (4 + 0 \cdot 4) \mod 7 = 4$$

0	~~Sz~~F	45
1	~~Sz~~F	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	81
5	Sz	
6	~~Sz~~F	20

Töröljük a $k = 127$-et.

$$127T \quad \left.\begin{aligned} h_0(127) &= 127 \mod 7 = 1 \\ h_1(127) &= 1 + (127 \mod 6) = 2 \end{aligned}\right\} \quad \Rightarrow$$

$$h(127, 0) = (1 + 0 \cdot 2) \mod 7 = 1$$

0	~~Sz~~F	45
1	~~SzF~~T	127
2	Sz	
3	~~SzF~~T	17
4	~~Sz~~F	81
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 3$-at.

$$\left.\begin{aligned} h_0(3) &= 3 \mod 7 = 3 \\ h_1(3) &= 1 + (3 \mod 6) = 4 \end{aligned}\right\} \quad \Rightarrow$$

$$h(3,0) = (3 + 0 \cdot 4) \mod 7 = 3$$

0	~~Sz~~F	45
1	~~SzF~~T	127
2	Sz	
3	~~SzFT~~F	~~17~~3
4	~~Sz~~F	81
5	Sz	
6	~~Sz~~F	20

Beszúrjuk a $k = 10$-et.

$$\left.\begin{aligned} h_0(10) &= 10 \mod 7 = 3 \\ h_1(10) &= 1 + (10 \mod 6) = 5 \end{aligned}\right\} \quad \Rightarrow$$

$$\begin{aligned} h(10,0) &= (3 + 0 \cdot 5) \mod 7 = 3 \\ h(10,1) &= (3 + 1 \cdot 5) \mod 7 = 1 \end{aligned}$$

0	~~Sz~~ F	45
1	~~Sz~~ ~~F~~ ~~T~~ F	~~127~~ 10
2	Sz	
3	~~Sz~~ ~~F~~ ~~T~~ F	~~17~~ 3
4	~~Sz~~ F	81
5	Sz	
6	~~Sz~~ F	20

A FELOSZT algoritmus

Az algoritmus feladata

Egy A tömb $A[p \ldots r]$ résztömbjét egy adott $x \in \mathbb{R}$ szám körül felosztja az $A[p \ldots q]$, $A[q \ldots r]$ két résztömbre olyan módon, hogy

- az $A[p \ldots q]$ résztömbbe az A tömbnek azok az elemei kerülnek, amelyek kisebbek vagy egyenlőek x-nél,
- az $A[q + 1 \ldots r]$ résztömbbe azok az elemek kerülnek, amelyek nagyobbak vagy egyenlőek x-nél.

A pszeudokód

Feloszt algoritmus: FELOSZT(A, p, r, x, q)
INPUT: A a tömb, amelynek az $A[p \dots r]$ résztömbjét felosztjuk, $x \in \mathbb{R}$ az előre megadott érték, amely a felosztást szabályozza.
OUTPUT: A a megváltozott tömb, q a felosztás határa, $A[p \dots q]$, $A[q+1 \dots r]$ a kapott résztömbök.

FELOSZT(A, p, r, x, q)
$i \longleftarrow p-1$, $j \longleftarrow r+1$
WHILE IGAZ DO
REPEAT $j \longleftarrow j-1$
UNTIL $A[j] \leqslant x$
REPEAT $i \longleftarrow i+1$
UNTIL $x \leqslant A[i]$
IF $i < j$
THEN CSERE$(A[i], A[j])$
ELSE $q \longleftarrow j$, RETURN(A, q)

A feloszt algoritmus működése

- Gondolatban két nyilat mozgatunk. A nyilak indexeket jelölnek, a nyilak a vizsgált indexű elemre mutatnak. A bal oldali nyilacska az i (index), a jobb oldali a j. A nyilakat kiinduló helyzetben tartományon kívülre, a tartomány szélére állítjuk. Mindkét nyilat befelé mozgatjuk.
- A j indexet jelölő nyilacskát mindaddig mozgatjuk balra, (azaz csökkentjük), amíg az elem amire mutat nagyobb x-nél. Amint olyan elemre mutat, ami egyenlő x-el, vagy kisebb nála, megállunk.
- Az i indexet jelölő nyilacskát mindaddig mozgatjuk jobbra, (azaz növeljük), amíg az elem amire mutat kisebb x-nél. Amint olyan elemre mutat, ami egyenlő x-szel, vagy nagyobb nála, megállunk.

- Ekkor megvizsgáljuk a két indexet. Ha i kisebb j-nél, akkor megcseréljük azokat az elemeket, amelyekre az i illetve a j mutat és tovább mozgatjuk befelé a nyilacskákat. Ha azonban az i és a j átfedik egymást (i=j), vagy j megelőzi az i-t (j<i), akkor leáll az algoritmus, a q megkapja a j értékét és visszatérünk q-val és a megváltozott tömbbel. A q jelöli a felosztás határát.

Érdemes megtanulni a feloszt algoritmust, hiszen ez működteti a gyors rendezést, ami fontos, illetve a kiválasztás lineáris időben algoritmust, ami érdekes.

Példa

Legyen $A = [17, 127, 20, 45, 81, 3, 10, 49, 50, 28, 12, 81, 9]$. Osszuk fel A-t az $x = 20$ körül.

	17	127	20	45	81	3	10	49	50	28	12	81	9	
$\uparrow$														$\uparrow$
		$\uparrow$											$\uparrow$	
		9	20	45	81	3	10	49	50	28	12	81	127	
			$\uparrow$								$\uparrow$			
			12	45	81	3	10	49	50	28	20			
				$\uparrow$			$\uparrow$							
				10	81	3	45							
					$\uparrow$	$\uparrow$								
					3	81								
					$\uparrow j$	$\uparrow i$								

$q \longleftarrow j = 5$, RETURN$(A, 5)$,
$A = [17, 9, 12, 10, 3 \mid 81, 45, 49, 50, 28, 20, 81, 127]$

Kiválasztási probléma

Legyen adott egy A halmaz (n db páronként különböző szám). Adott továbbá $1 \leqslant i \leqslant n$ pozitív egész. Keressük az A halmaznak azt az x elemét, amelynél pontosan $i - 1$ db kisebb eleme van a halmaznak.

Kiválasztás lineáris időben

Alsó medián

Medián páratlan elemszámú A halmaz esetén a rendezett A halmaz középső eleme, míg páros elemszám esetén a két középső közül a kisebb vagy egyenlő (alsó medián).
Azaz, ha $A = \{x_1, \ldots x_n\}$ egy halmaz, akkor először rendezzük a halmaz elemeit,

$$x_1^* \leqslant x_2^* \leqslant \cdots \leqslant x_n^*$$

ekkor

$$\mathrm{med} = x^*_{\lceil \frac{n}{2} \rceil}.$$

KÖSZÖNÖM A FIGYELMET!