

ZADAĆA 1 - HASH TABLICE

Strukture podataka i algoritmi II - IO54

2.3.2023.

Zadatak 1. (5 + 5 + 10) Sljedeći zadatak bavi se ilustracijom korištenja funkcije raspršivanja (engl. hash functions).

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 rješavaju ulančavanjem, gdje je dana funkcija raspršenja $h(k) = k \bmod m$.
 - b. u kojem se kolizije rješavaju probiranjem za $i = 0, 1, 2, \dots, m - 1$, koristeći dvostruko probiranje

$$h(k, i) = (h_1(k) + i * h_2(k)) \bmod m$$

gdje su $h_1(k) = k \bmod m$, $h_2(k) = 1 + (k \bmod (m - 1))$ pomoćne hash funkcije.

2. Promotrite n -znamenkasti decimalni broj $x_1x_2\dots x_n$ ($x_i \in \{0, 1, \dots, 9\}$). Je li hash funkcija $f(x) = \sum_{i=1}^n a_i x_i \bmod 8$ univerzalna? Za a_i , $i = 1, 2, \dots, n$ nezavisne slučajne varijable uniformno izabrane iz $\{0, 1, 2, 3, 4, 5, 6, 7\}$? Obrazložite, ako je dajte razlog zašto je, ako nije dajte kontraprimjer.

Zadatak 2. (15) Pretpostavimo da koristimo hash funkciju h da se rasprši n različitih ključeva u tablicu T duljine m . Uz pretpostavku uniformnog raspršenja, koliki je očekivani broj kolizija? Preciznije, koliki je očekivani kardinalitet $\{ \{k, l\} : k \neq l, h(k) = h(l) \}$?

Zadatak 3. (5 + 5 + 10 + 15) Tablica raspršenja (hash tablica) veličine m koristi se za spremanje n ključeva, gdje $n \leq m/2$. Neka je korišteno otvoreno adresiranje s probiranjem za rezolviranje kolizija:

1. Uz pretpostavku uniformnog raspršenja, pokažite da za $i = 1, 2, \dots, n$ vjerojatnost da i -to ubacivanje zahtjeva strogo više od k probiranja najviše 2^{-k} .

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

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

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

4. Pokažite da je očekivana duljina $E[X]$ najdužeg niza probiranja $O(\lg n)$.

Zadatak 4. (10 + 10 + 10) Implementirajte strukturu podataka rječnika koja se sastoji od kolekcije jedinstvenih parova (k, value) gdje k predstavlja ključ, a value pridruženu vrijednost. Koristite tablice raspršenja za implementaciju takve strukture, za rješavanje kolizija koristite funkciju iz Zadatka 1. pod b). Možete pretpostaviti da je k prirodan broj. Struktura mora imati sljedeće funkcionalnosti:

- a. pretraživanje: za dani ključ k pronaći (k, value) u strukturi ili vratiti NIL.
- b. dodavanje: omogućiti dodavanje para (k, value) u strukturu. Ukoliko postoji ključ k u strukturi onda treba ažurirati value u paru (k, value) .
- c. brisanje: za dani ključ k obrisati par (k, value) u strukturi.

Tablica u strukturi neka bude vektor parova $(\text{int}, \text{float})$. Sve ključeve iz zadatka 1. ubacujte redom kako su napisani u zadatku (**bitno**), tako da value za prvog bude 0.1, za drugog 0.2, itd.

Ispis neka bude u obliku (primjer sa drugim keyevima):

K:33 V:0.5 K:1 V:0.7 K:11 V:0.6 K:3 V:0.2 K:15 V:0.1 K:2 V:0.8 K:10 V:0.9 nil nil K:20 V:0.3 K:21 V:0.4 gdje jedan nil predstavlja jedan element a ne K i V.

Prvo ubacite sve elemente iz zadatka 1. i ispišite tablicu.

Zatim obrišite par s ključem 40 i ažurirajte value na ključu 39 na 1.1, te zatim ponovo ispišite cijelu tablicu.