

Algoritmusok gyakorlat

9. gyakorlat, november 2.

9. gyakorlat

Szótár: egy halmaz elemeihez (kulcs) egy érték adat tartozik

Bemenet: egy szótár és egy keresett kulcs

Kimenet:

- Ha a *kulcs* szerepel a szótárban akkor a hozzá tartozó *érték*
- Ha nem szerepel *kulcs* a szótárban akkor NIL

9. gyakorlat

Új adatszerkezetek

Halmaz, csak az számít, hogy tartalmaz-e egy elemet vagy sem:

halmaz (Set)

Ha kulcs-érték párokat tárolunk:

szótár (dictionary, Map)

9. gyakorlat

Csak a Beszúr, Keres és Töröl szótárműveleteket támogató dinamikus halmazra van szükségünk! A hasító táblázat hatékony adatszerkezet szótárak megvalósítására. A hasító táblázat a megszokott tömbfogalom általánosítása.

Közvetlen címzésű táblázat: kis méretű kulcs univerzum van !

Ha a szótárban tárolt kulcsok halmaza (n) sokkal kisebb a lehetséges kulcsok univerzumánál – hasító táblázatok alkalmazása.

9. gyakorlat

h hasító függvény a kulcsok U univerzumát képezi le a $T[0 \dots m - 1]$ ***hasító táblázat*** ***réseire***: $h : U \rightarrow \{0, 1, \dots, m-1\}$. A tömb elemeinek közvetlen címzésére a kulcs helyett egy, a kulcsból *kiszámított* érték szolgál. A "láncolás" az "ütközések" – amikor is a hasító függvény egynél több kulcsot képez le ugyanarra a tömbindexre – egyik kezelési módszere. ***A láncolásnál*** az ugyanarra a résre leképeződő elemeket összefogjuk egy láncolt listába.

9. gyakorlat

A hasító függvények megadására szolgáló **osztásos módszer** esetén egy k kulcsot úgy képezünk le az m rés valamelyikére, hogy vesszük k m -mel való osztásának maradékát. Azaz a hasító függvény a következő: $h(k) = k \bmod m$. Ha például a hasító táblázat mérete $m = 12$ és a kulcs $k = 100$, akkor $h(k) = ?$

A szorzásos módszer a következő: $h(k) = \lfloor m(kA \bmod 1) \rfloor$, ahol $kA \bmod 1$ a kA törtrészét jelenti. Pl. $m=128$, $A=0.1$, $k=48$.

9. gyakorlat

Kövessük végig láncolósos ütközésfeloldás esetén az 5, 28, 19, 15, 20, 33, 12, 17, 10 kulcsok hasító táblázatba való beszúrását. Legyen a rések száma 9 és a hasító függvény $h(k) = k \bmod 9$.

9. gyakorlat

Alapfogalmak a hatékonysághoz

m méretű T tömb, n elem

kitöltési tényező $\alpha = n/m$ (=láncok átlagos hossza)

bármely j -re a $T[j]$ lista hosszának várható értéke

$$E[n_j] = n/m = \alpha$$

$h(k)$ hasított érték $O(1)$ idő alatt számítható ki

9. gyakorlat

Ütközésfeloldás nyílt címzéssel

Minden vödörbe max egy kulcsot teszünk. A hasító kód alapján számolt index csak kiindulópont. Ha foglalt, akkor valamilyen módszerrel végigpróbálunk másokat. Jelölés: $\text{proba}(h(k), i)$: a $h(k)$ kulcs i . próbája

- lineáris: $\text{proba}(h(k), i) = (h(k) + i) \bmod N$
- négyzetes: $\text{proba}(h(k), i) = (h(k) + c_1 \cdot i + c_2 \cdot i^2) \bmod N$
- dupla hasítás: $\text{proba}(h(k), i) = (h_1(k) + i \cdot h_2(k)) \bmod N$

Ha a kitöltési tényező kb. 0.7-re növekszik, mindegyik módszer meglassul.

11. gyakorlat

Szűrjük be egy 10 hosszú hasító táblába a következő objektumokat a láncolás, lineáris, négyzetes ($c1 = 0$, $c2 = 1$) és dupla hasítás módszerekkel. Az elemek két hasító függvény ($h1, h2$) által adott hasító kódjai a következők:

$h1(\text{"cím"}) = 210$, $h2(\text{"cím"}) = 567$

$h1(\text{"foo"}) = 130$, $h2(\text{"foo"}) = 132$

$h1(\text{"bar"}) = 65$, $h2(\text{"bar"}) = 324$

$h1(\text{"fák"}) = 188$, $h2(\text{"fák"}) = 111$

$h1(\text{"bin"}) = 785$, $h2(\text{"bin"}) = 176$

$h1(\text{"one"}) = 960$, $h2(\text{"one"}) = 608$

9. gyakorlat

A nyílt címezéses hasító táblázatokból való törlés kissé nehezebb. Amikor töröljük az i -edik részben lévő kulcsot, nem elegendő a nil beírásával jelezni annak ürességét, mert ha csak ezt tennénk, akkor lehetetlen lenne minden olyan kulcs visszakeresése, melynek beszúrása során az i -edik részt kipróbáltuk és foglaltnak találtuk. Egy lehetséges megoldás az, hogy a törlés jelölésére a nil helyett egy másik speciális értéket, a törölt-et használjuk.

9. gyakorlat

Adott egy nyílt címzéses, $m = 11$ hosszúságú (üres) hasító táblázat a $h(k) = k \bmod m$ elsődleges hasító függvénnnyel, s tekintsük a beszúrandó 10, 22, 31, 4, 15, 28, 17, 88, 59 kulcsokat. Szemléltessük a beszúrás eredményét akkor, ha a lineáris kipróbálást, a $c_1 = 1$ és $c_2 = 3$ állandókat használó négyzetes kipróbálást, illetve a $h_2(k) = 1 + (k \bmod (m - 1))$ másodlagos hasító függvényű dupla hasítási módszert alkalmazzuk.