

Datové struktury

2. ledna 2013

Obsah

1	Hashování	2
1.1	Hashování se separovanými řetězci	2
1.2	Hashování s přemísťováním	2
1.3	Hashování s dvěma ukazateli	2
1.4	Srůstající hashování	3

1 Hashování

Chceme implementovat množinu s operacemi:

- Insert(x)
- Delete(x)
- Member(x)

Pomocí bitového pole to jde zvládnout rychle, ale když je velké univerzum, je to pole velké.

Proto vezmeme nějakou funkci, která vezme prvek a vrátí rozumě malé číslo, používá se toto číslo místo toho. Ovšem vznikají kolize (více prvků patří na stejné místo), je třeba řešit.

1.1 Hashování se separovanými řetězci

V každé pozici bydlí spoják všech věcí, které bydlí v tomto místě. Potom je potřeba spojáček prohledávat.

Poté je potřeba nějak zjišťovat rychlost. První možnost je vše spočítat, je to pracné, ale přesné. Druhá možnost je měřit to, ale je to nepřesné, je potřeba dělat statistiky. V tomto případě, nejhorší možnost je, že vše se to nacpe do jednoho řetězce, nejhorší složitost je $O(n)$.

Pokud máme nějaké předpoklady, například že S je náhodně vybraná z universa a že h rozkládá prvky náhodně. Nemusí platit. A v praxi často neplatí – poté se děly špatné věci.

Uvažujeme faktor naplnění $\alpha = \frac{n}{m}$. Je třeba spočítat očekávanou průměrnou a maximální délku řetězce, průměrný počet testů pro úspěšné resp. neúspěšné vyhledání.

1.2 Hashování s přemísťováním

Podobně, jako u separovaných řetězců, ale položky jsou uvnitř tabulky a odkaz je index do tabulky. Spotřebává méně paměti, je lepší ke cachím.

1.3 Hashování s dvěma ukazateli

Podobně jako u separovaných řetězců, jednosměrné seznamy. Řetězec může začít kdekoliv, na to je položka begin – když řetězec začíná někde jinde (položky jsou sice přímo v hashovací tabulce, ale pozice zcela nesouvisí s hodnotou).

1.4 Srůstající hashování

Buď standardní a nebo s pomocnou pamětí. Prvky se strkají do tabulky a pokračují za sebou. Když není volno, pokračujeme na další a další, dokud volno nebude. Tím se mohou spojit některé řetězce dohromady. Dělí se jednak, jestli použijeme pomocnou paměť (aby nesrůstaly) a jednak podle toho, kam se vkládá nový prvek v případě kolize.