

Funzioni hash

La funzione hash da usare per una tabella dovrebbe essere scelta in modo da garantire quanto più possibile la proprietà di simple uniform hashing.

Metodo della divisione

$$h(k) = k \bmod m$$

In questo caso è preferibile che $m \neq 2^p$, in quanto in tal caso l'operazione equivarrebbe ad estrarre i p bit meno significativi della chiave; se la chiave è rappresentazione binaria di un oggetto, questo porterebbe a ignorare alcuni dei suoi campi. Se invece m è un numero primo lontano da una potenza di 2 i bit presi in considerazione saranno distribuiti più equamente.

Metodo della moltiplicazione

$$h(k) = \lfloor m(kA \bmod 1) \rfloor$$

con $0 < A < 1$ e $x \bmod 1 = x - \lfloor x \rfloor$. In questo possiamo scegliere $m = 2^p$ e $A = \frac{s}{2^w}$ (con w la dimensione di una parola e $0 < s < 2^w$) e calcolare efficientemente l'hash moltiplicando k (w bit) per s : il risultato è un numero da $2w$ bit con valore $r_1 2^w + r_0$, possiamo quindi osservare i p bit più significativi di r_0 (la parola meno significativa) per ottenere il valore di $h(k)$.