

Appunti lezione – Capitolo 7

Tabelle Hash

Alberto Montresor

19 Agosto, 2014

1 Domanda - Perché $2^p - 1$ non va bene per il metodo della divisione

Data una sequenza di caratteri presi da un alfabeto di dimensione 2^p , se si utilizza un valore di m pari a $2^p - 1$, due sequenze permutate danno origine allo stesso valore hash.

Dimostrazione

$$\begin{aligned} & [2^p a + b] \bmod 2^p - 1 \\ = & [(2^p - 1)a + a + b] \bmod 2^p - 1 \\ = & [a + b] \bmod 2^p - 1 \\ = & [(2^p - 1)b + b + a] \bmod 2^p - 1 \\ = & [2^p b + a] \bmod 2^p - 1 \end{aligned}$$

2 Domanda: Costo computazionale di una ricerca senza successo, in caso di indirizzamento aperto

Il numero atteso di ispezioni $I(\alpha)$ per una ricerca senza successo è pari a $\frac{1}{(1-\alpha)}$.

Abbiamo un tabella di m celle, con n elementi già inseriti. Quindi $n/m = \alpha$.

In un ricerca senza successo, ogni ispezione, tranne l'ultima, trova una cella occupata; l'ultima trova una cella **nil**.

- Variabile casuale X : numero di ispezioni.
- Evento A_i : l' i -esima ispezione trova una cella occupata.

$$\begin{aligned} Pr(X \geq i) &= Pr(A_1 \wedge A_2 \wedge \dots \wedge A_{i-1}) \\ &= Pr(A_1) \cdot Pr(A_2|A_1) \cdot Pr(A_3|A_1 \wedge A_2) \cdot \dots \cdot Pr(A_{i-1}|A_1 \wedge \dots \wedge A_{i-2}) \\ &= \frac{n}{m} \cdot \frac{n-1}{m-1} \cdot \frac{n-2}{m-2} \cdot \dots \cdot \frac{n-(i-2)}{m-(i-2)} \\ &\leq \left(\frac{n}{m}\right)^{i-1} \\ &\leq \alpha^{i-1} \end{aligned}$$

Questo perché alla prima ispezione, ci sono n elementi in m celle; alla seconda ispezione, non si può più passare per la prima cella, quindi ci sono $n - 1$ elementi in $m - 1$ celle; e così via.

$$\begin{aligned}
 E[X] &= \sum_{i=1}^{\infty} Pr(X \geq i) \\
 &\leq \sum_{i=1}^{\infty} \alpha^{i-1} \\
 &= \sum_{i=0}^{\infty} \alpha^i \\
 &= \frac{1}{1 - \alpha}
 \end{aligned}$$

3 Domanda: Costo computazionale di una ricerca con successo, in caso di indirizzamento aperto

Una ricerca con successo per la chiave k segue la stessa sequenza di quando k è stato inserito.

Sia β il fattore di carico all'atto dell'inserzione ed α quello all'atto della ricerca con successo. Facendo la media sui possibili valori di β , che sono compresi tra 0 ed α , si ottiene:

$$S(\alpha) = \frac{1}{\alpha} \int_0^{\alpha} \frac{1}{1 - \beta} d\beta = \frac{1}{\alpha} \ln \frac{1}{1 - \alpha} = -\frac{1}{\alpha} \ln(1 - \alpha)$$