

Hash Table

Pandolfini Luca

September 2020

1 Introduzione

In questa relazione andro' a studiare il comportamento delle tabelle hash al crescere del fattore di caricamento.

$$\alpha = n/m. \tag{1}$$

Dove n rappresenta il numero di valori inseriti, e m il numero delle chiavi a cui associarli.

2 Struttura

Ho implementato la tabella hash con gestione delle collisioni basate su concatenamento e su indirizzamento aperto. La funzione hash, ad ispezione lineare, e' calcolata col metodo delle divisioni.

3 Test

I test che andro' ad eseguire non saranno altro che degli inserimenti variabili in tabella, registrando il numero di collisioni effettuate in funzione del fattore di caricamento, sia in indirizzamento aperto (plot rosso), sia con concatenamento (plot: blue), mi aspetto che la tabella ad indirizzamento aperto all'aumentare del fattore di carico aumenti maggiormente il numero di collisioni rispetto a quella con concatenamento, in quanto per gestire le collisioni riempie le chiavi disponibili piu' velocemente, aumentando la probabilita' di ulteriori collisioni.

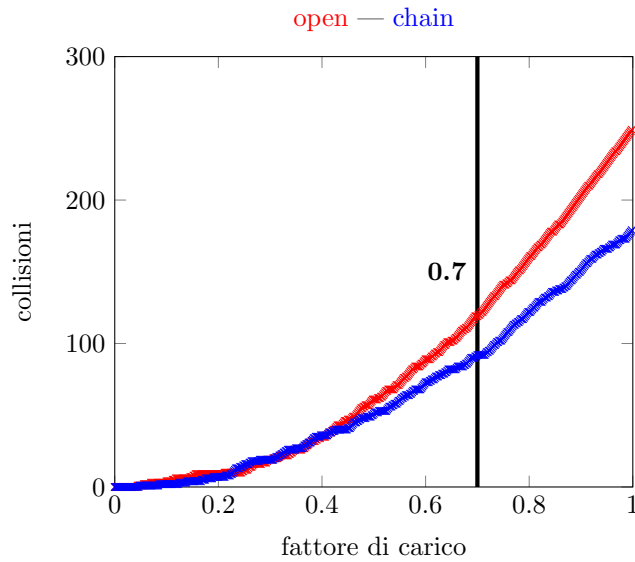


Figure 1: 500 keys, riempite con valori random

4 Conclusioni

Come atteso, la curva delle collisioni col metodo di concatenamento e' piu bassa rispetto a quella dell' indirizzamento aperto.

In nero e' rappresentato il coefficiente ottimale del fattore di carico, che si aggira intorno al 0.7, la differenza tra le due gestioni di collisioni e': per fattori di carico basso e' migliore quella ad indirizzamento aperto in quanto effettua ricerche piu' veloci perche' i valori sono posizionati in memoria in modo contiguo, mentre col concatenamento la ricerca viene effettuata su liste con puntatori, mentre per fattori di carico maggiori di 1 per la open serve un re-hashing mentre col concatenamento si puo' mantenere lo stesso numero di chiavi.