

Ayudantía 8

Carlos Lagos - carlos.lagosc@usm.cl

¿Qué es un Diccionario?

Un diccionario es una estructura de datos que almacena valores de manera que cada valor está asociado a una llave única. En un diccionario, cada llave puede estar relacionada con un solo valor.

¿Qué es un Diccionario Ordenado?

Un diccionario ordenado es una variante de diccionario en la que las llaves están organizadas siguiendo un orden específico. Podemos modelar un diccionario ordenado utilizando un árbol binario, donde cada nodo almacena una llave y su valor asociado.

¿Qué es un Diccionario Desordenado?

Un diccionario desordenado es una estructura de datos en la que no se requiere un orden específico para las llaves. Un ejemplo de esto sería almacenar los pares de llave-valor en una lista de tuplas. Sin embargo, la búsqueda de un valor en una lista de tuplas no es tan eficiente.

¿Qué es una Función de Hashing?

Una función de hashing es una función que recibe un elemento y produce otro. En este caso, normalmente nos conviene que al ingresar un elemento se devuelva un número. Esto se debe a que comúnmente resulta útil relacionar un elemento con un número.

¿Qué es una Tabla de Hash?

Una tabla de hash utiliza una función de hash para relacionar una llave con una posición en un arreglo. De esta forma, podemos guardar el valor asignado a la llave en esa posición específica del arreglo. Comúnmente, para generar un número que corresponda a una posición válida en el arreglo, se usa la operación módulo.

Manejo de Colisiones en Tablas de Hash

Debido al espacio limitado, es posible que dos llaves generen el mismo valor de hash. Para resolver esto, existen dos enfoques:

- Hashing Cerrado: Si ocurre una colisión, se busca otra posición disponible en la tabla.
- Hashing Abierto: Cada posición de la tabla utiliza una estructura de datos, como una lista enlazada o un árbol binario. Si hay una colisión, el nuevo elemento se añade a la estructura de datos correspondiente en esa posición.

Políticas de Colisiones Comunes

Algunas políticas de colisiones comunes en el hashing cerrado incluyen:

- Sondeo Lineal
- Sondeo Cuadrático
- Hashing Doble

Consideraciones al Borrar un Registro en una Tabla de Hashing

Hay dos consideraciones importantes al borrar un registro en una tabla de hashing:

- Mantener la Continuidad de Búsqueda
- Estados de las Ranuras

Una ranura liberada debe poder ser utilizada en futuras inserciones.

Hashing Cerrado: Factor de Carga

El factor de carga en una estrategia de hashing cerrado se define como:

$$\alpha = \frac{N}{M}$$

Donde N es la cantidad actual de registros en la tabla y M es la capacidad total de la tabla.

- Indica el porcentaje de llenado de la tabla.
- Un factor de carga de $\alpha = 0.6$ (60% de llenado) suele ser un buen balance.

Probabilidad de i colisiones

Asumamos:

- Que la secuencia de ranuras generada ante colisiones es una permutación aleatoria de las ranuras de la tabla.
- Que todas las ranuras de la tabla tienen la misma probabilidad de ser la ranura original para el siguiente registro a insertarse.

La probabilidad de i colisiones:

$$\frac{N(N-1)\dots N(N-i+1)}{M(M-1)\dots M(M-i+1)} \approx \left(\frac{N}{M}\right)^i = \alpha^i$$

Tablas de Hash

Supongamos que tenemos la siguiente secuencia de llaves y valores:

- Valores (llave-valor): [(12, 'A'), (25, 'B'), (38, 'C'), (47, 'D'), (56, 'E'), (69, 'F'), (73, 'G'), (84, 'H'), (95, 'I')]

Se tiene un arreglo de 17 espacios disponibles.

Dada la siguiente función de hashing y política de colisiones:

- Función de Hashing: $h(k) = (2k + 5) \mod 17$
- Política de Colisiones: $p(k, i) = 3i \mod 17$

Tablas de Hash

Supongamos que tenemos la siguiente secuencia de llaves y valores:

- Valores (llave-valor): [(20, 'X'), (35, 'Y'), (42, 'Z'), (53, 'W'), (67, 'U'), (72, 'V'), (89, 'T'), (94, 'S'), (108, 'R')]

Se dispone de un arreglo de 17 espacios disponibles.

Dada la siguiente función de hashing y política de colisiones:

- Función de Hashing: $h(k) = (3k^2 + 7) \mod 17$
- Política de Colisiones: $p(k, i) = (2i^2 + 5i) \mod 17$