



Universidad
Rafael Landívar

Tradición Jesuita en Guatemala

Tablas Hash

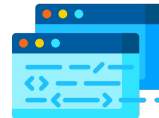
Búsqueda y almacenamiento eficiente de datos

Introducción

Las tablas hash son estructuras de datos eficientes para **almacenar** y **buscar** pares clave-valor.

Funcionan mediante una **función hash** que asigna una clave a una ubicación en la tabla.

Las tablas hash son más útiles cuando se almacenan grandes cantidades de información.



Tablas Hash

Para usar una tabla hash se necesita:

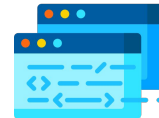
1. Una estructura de acceso directo (normalmente un array)
2. Una estructura de datos con una clave o un elemento que funja como clave
3. Una función resumen (hash) cuyo dominio sea el espacio de claves y su rango los números naturales



Terminología

Clave: Elemento individual o propiedad de un objeto que servirá como entrada para la función hash y así obtener el valor hash.

Valor: Resultado de la función hash que nos indicará el índice adecuado para inserción y búsqueda en la tabla.

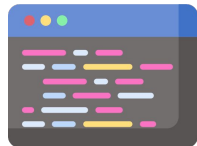


Función Hash

En esencia la función hash toma una clave y la **convierte en un índice** dentro de la tabla.

Debe ser rápida y generar índices uniformemente distribuidos.

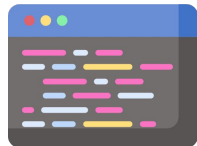
Para almacenar un elemento en la tabla hash se ha de convertir su clave a un número.



Eficiencia

Las tablas hash se suelen implementar sobre vectores de una dimensión, aunque se pueden hacer implementaciones multi-dimensionales basadas en varias claves.

Como en el caso de los arrays, las tablas hash **proveen tiempo constante de búsqueda promedio $O(1)$** , sin embargo, en escenarios no favorables el tiempo de búsqueda puede llegar a $O(n)$.

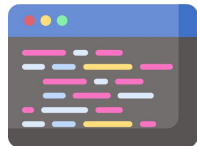


Inserción

Para este método es sumamente importante aclarar que el algoritmo detrás de la función hash **debe considerar con la misma probabilidad todos los índices** para insertar. No pueden haber preferencias sobre ciertos índices. Todos los slots deben ser considerados para la inserción.

Ejemplo de una función hash simple:

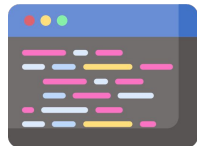
$hash(clave) = clave \% tamaño_tabla$



Búsqueda

La forma de implementar en función esta operación es pidiendo la clave y con esta **devolver el valor hash**.

1. Para recuperar los datos, es necesario únicamente conocer la clave del elemento, a la cual se le aplica la función hash.
2. El valor obtenido se mapea al espacio de direcciones de la tabla.
3. Si el elemento existente en la posición indicada en el paso anterior tiene la misma clave que la empleada en la búsqueda, entonces es el deseado.



Aplicaciones

- Las tablas hash se utilizan con frecuencia para indexar y buscar volúmenes masivos de datos. Un motor de búsqueda puede utilizar una tabla hash para almacenar las páginas web que ha indexado.
- Los datos generalmente se almacenan en la memoria caché a través de tablas hash, lo que permite un acceso rápido a la información de uso frecuente.



Aplicaciones

- Las funciones hash se utilizan con frecuencia en criptografía para crear firmas digitales, validar datos y garantizar la integridad de los datos.
- Las tablas hash se pueden utilizar para implementar índices de bases de datos, lo que permite un acceso rápido a los datos basados en valores clave.



Colisiones

¿Qué sucede si dos elementos/objetos **producen un mismo valor hash**?

La respuesta es que se producen colisiones. Al momento de tratar con colisiones lo primero que hay que considerar es el tamaño de la tabla, entre más grande sea la tabla menor será la probabilidad de colisiones.





Universidad
Rafael Landívar

Tradición Jesuita en Guatemala

Tablas Hash

Búsqueda y almacenamiento eficiente de datos