



10

CLASE

ESTRUCTURA DE DATOS Y PROGRAMACIÓN METÓDICA

POR ANA RONCAL



EMPEZAR →



ALMACENAMIENTO HASH

En esta clase vamos a explorar el concepto de **funciones hash**, que significa que el orden (y más eficientemente la ubicación de un elemento dentro de una colección) está determinado por alguna función del valor del elemento que hay que almacenar, o por alguna función de un valor clave del elemento que hay que almacenar.



ALMACENAMIENTO HASH

Ana
David
Fabiola
Gaby
Hector
Karla

Con las **funciones hash**, los elementos se almacenan en una **tabla de hash**, estando su ubicación dentro de una tabla determinada por una **función hash**. Cada ubicación de la tabla se denomina celda o franja.



ALMACENAMIENTO HASH

Ana
David
Fabiola
Gaby
Hector
Karla

Suponga que tenemos un arreglo de 12 elementos donde queremos almacenar nombres, por lo que creamos una función hash que **asigne a cada nombre una posición dentro del arreglo**, que estará asociada con la primera letra del nombre.



ALMACENAMIENTO HASH

Ana
David
Fabiola
Gaby
Hector
Karla

Por ejemplo, la primera letra A se asignará a la posición 0 del arreglo. la primera letra D se asignará a la posición 3.


Observe que, el **tiempo de acceso** a un elemento concreto es **independiente** al número de elementos contenidos en la tabla.

ALMACENAMIENTO HASH

Ya no hace falta realizar comparaciones para encontrar un elemento concreto o para averiguar la posición apropiada para un cierto elemento. **Con el almacenamiento hash, simplemente calculamos donde debe encontrarse cada elemento en particular.**

Sin embargo, esta eficiencia sólo puede conseguirse completamente si cada elemento se asigna a una posición distintiva dentro de la tabla.

ALMACENAMIENTO HASH



Ana
David
Fabiola
Gaby
Hector
Karla

¿Qué sucedería si tratáramos de almacenar el nombre "Ana" y el nombre "Andrés".

Esta situación en la que **dos elementos o claves se asignan a la misma ubicación de la tabla** se denomina **colisión**.

ALMACENAMIENTO HASH

Aunque resulta posible definir funciones **hash perfectas**, una función hash que se limite a distribuir adecuadamente los elementos entre las distintas posiciones de la tabla también proporcionará como resultado un tiempo de acceso constante $O(1)$ a los elementos de la tabla.



ALMACENAMIENTO HASH

PRESTA ATENCIÓN A:

Cómo debe ser de grande la tabla.

Si conocemos el tamaño del conjunto de datos, una regla es hacer el tamaño de datos unos 150 por ciento del tamaño del conjunto de datos.

Caso contrario, se utilizan técnicas de redireccionamiento dinámico. Se crea una tabla mayor que la original (doble), se pasan los datos.



ALMACENAMIENTO HASH

Una característica inherente a las tablas hash es que **su rendimiento se degrada significativamente a medida que se llenan.**

Se puede utilizar un **factor de carga**. El factor de carga de una tabla hash es el porcentaje de ocupación de la tabla para el cual deberá procederse a redimensionar está. Por ejemplo, si fijamos el factor de carga en 0.5, entonces la tabla se redimensionará cada vez que se alcance una ocupación del 0.5 por ciento.



FUNCIONES HASH

El método utilizado en el ejemplo anterior se denomina **extracción**. La extracción implica utilizar únicamente una parte de la clave con el fin de calcular la ubicación donde hay que almacenarlo. Otros métodos son:

El método de división,

El método de transformación de base

El método de análisis de dígitos

El método dependiente de la longitud

El método del cuadrado medio

El método de plegado, la clave se divide en una serie de partes



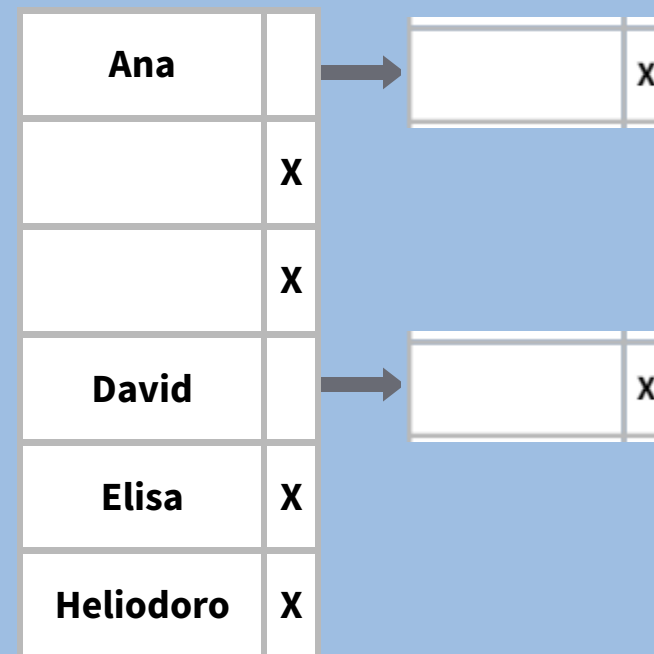
COLISIONES



RESOLUCIÓN DE COLISIONES

Encademamiento

Un método de encadenamiento para el tratamiento de colisiones trata la tabla como si fuera una tabla de colecciones. Cada celda de la tabla es un puntero a la coleccion asociada en dicha ubicación.





QUÉ APRENDIMOS HOY

- Con las **funciones hash**, los elementos se almacenan en una tabla de hash, estando su ubicación dentro de una tabla determinada por una función hash.
- El tiempo de acceso a un elemento concreto es **independiente** al número de elementos contenidos en la tabla.
- Dos elementos o claves se asignan a la misma ubicación de la tabla se denomina **colisión**.

BIBLIOGRAFÍA

Melgar, A. (2021). *Estructuras de Datos (TAD) y sus implementaciones. [PowerPoint slides]*. Facultad de Ciencias e Ingeniería. PUCP.

Cueva, R. (2022, Agosto 21). Capítulo 2. *Estructuras de Datos (TAD) y sus implementaciones. [PowerPoint slides]*. Facultad de Ciencias e Ingeniería. PUCP.

Hernández, R., Lázaro, J. C., Dormido, R. y Ros, S. (2001). *Estructuras de Datos y Algoritmos*. Pearson Education, S.A., Madrid

Lewis, J. y Chase, J. (2006) *Estructuras de datos con Java diseño de estructuras y algoritmos*. Pearson Education S.A., Madrid