

Hash

75.41 - Algoritmos y Programación II

1° Cuatrimestre 2021

Antes de empezar...



“Que es para ustedes un
Diccionario?”



Diccionario



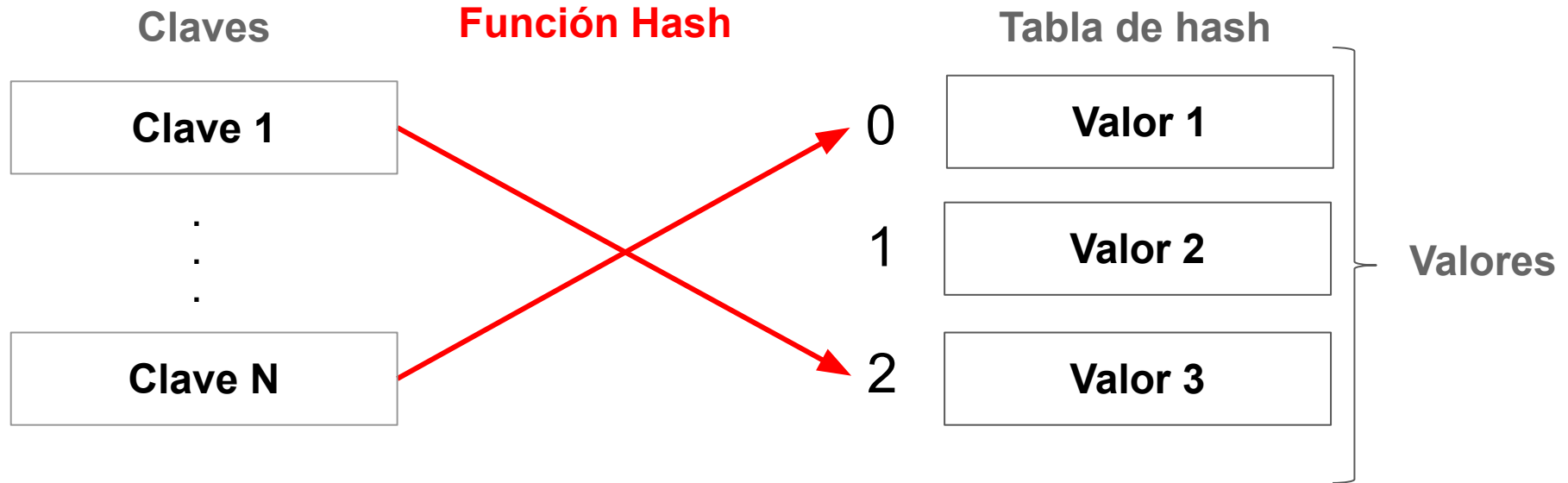
¿Porque usarlo?

- Performance a la hora de acceder al dato
- ¡No hay duplicación de entradas!

¿Cómo lo implementamos?

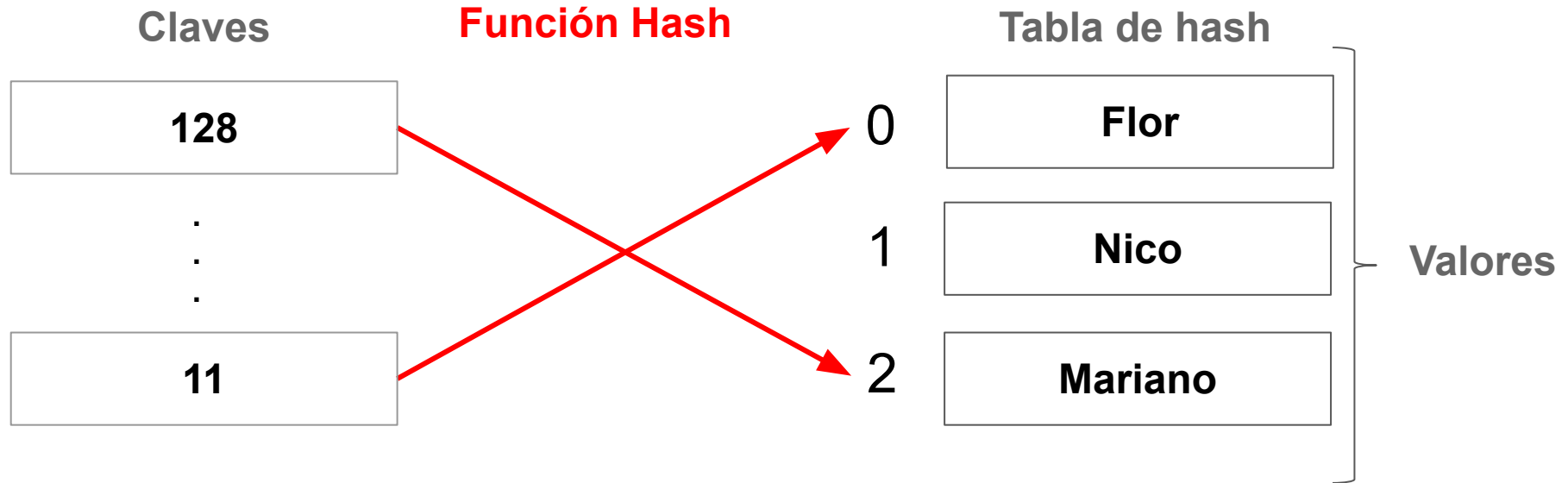
¿Qué es una tabla de hash?

- Estructura que contiene **valores**
- Puedo hallar un valor a partir de una **clave**



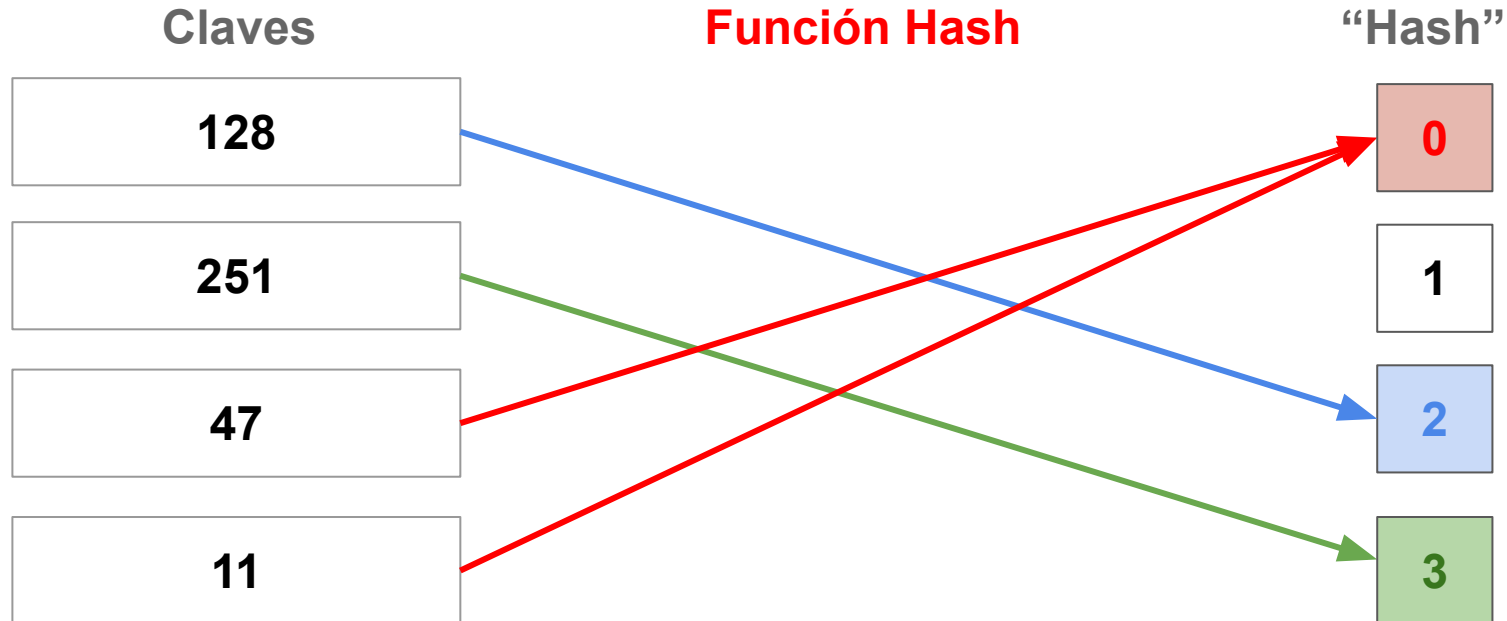
¿Qué es una tabla de hash?

- Estructura que contiene **valores**
- Puedo hallar un valor a partir de una **clave**



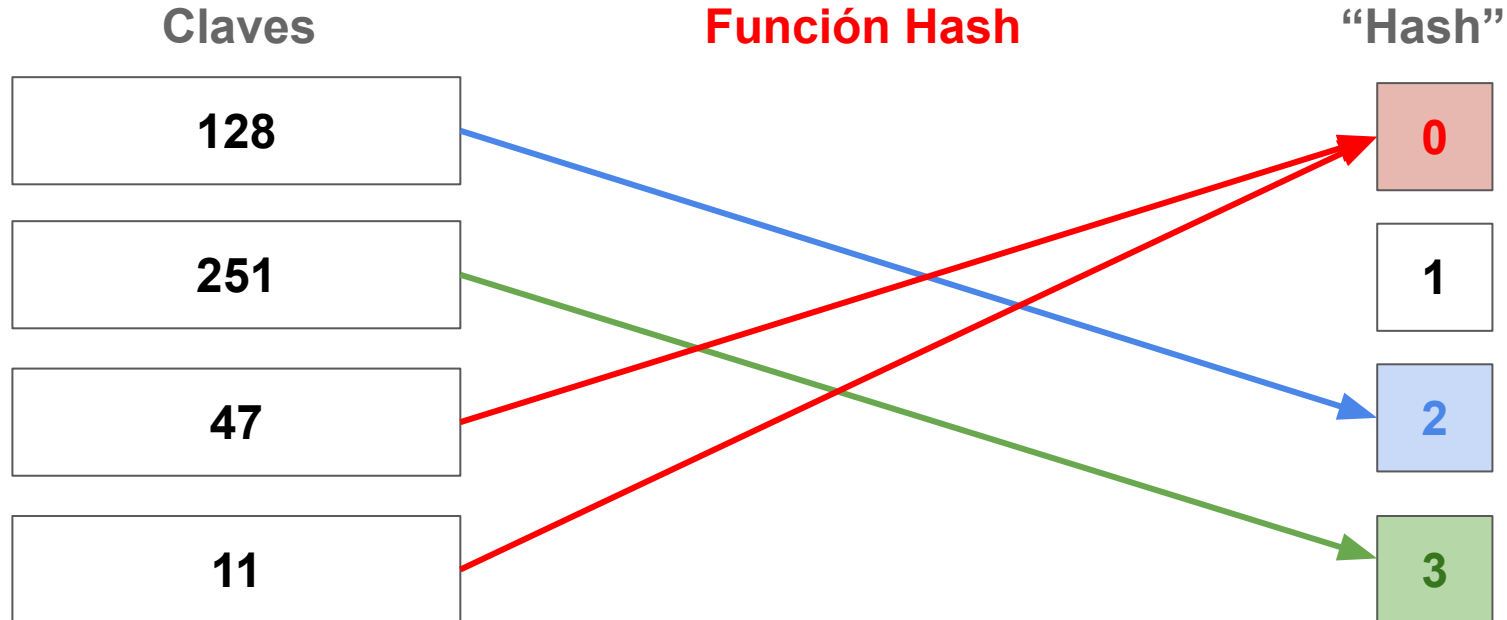
¿Qué es una función hash?

- Es una **función** que transforma claves en un número asociado



Colisiones

- Pueden haber más claves que espacios en la tabla de hash
 - Claves distintas dan el mismo valor de “hash”



Tipos de hash

**Direccionamiento
Cerrado**



**Hash
Abierto**



**Direccionamiento
Abierto**



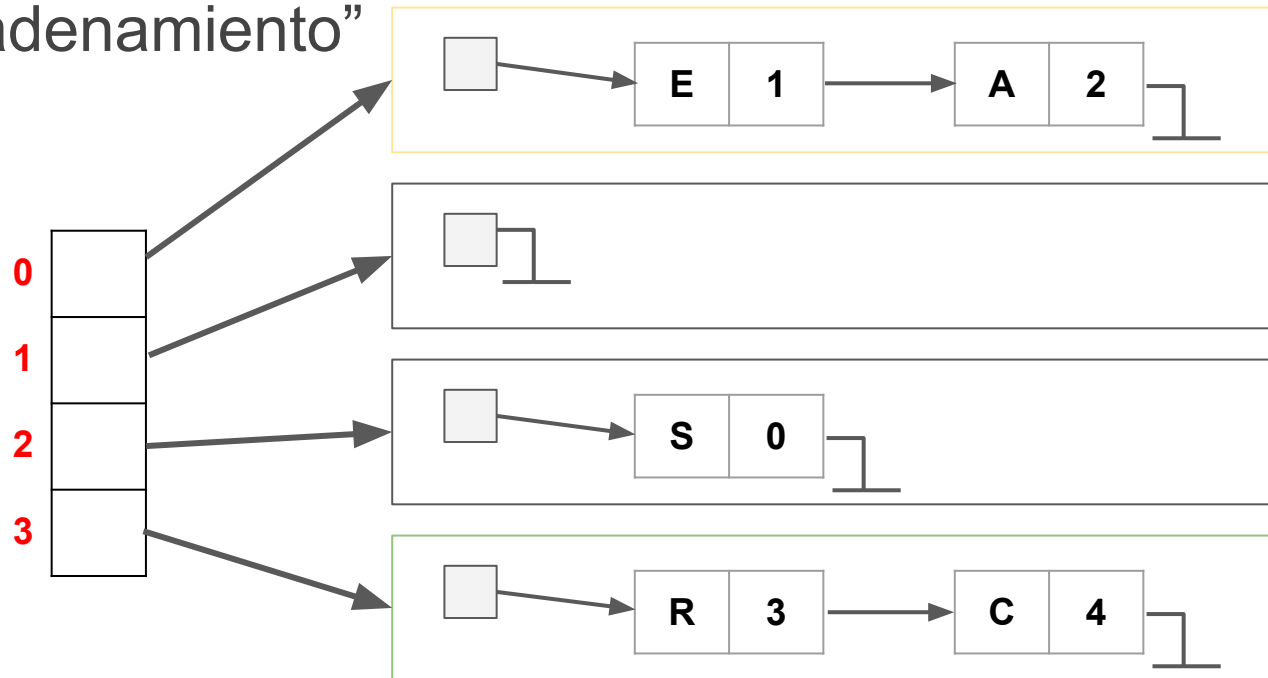
**Hash
Cerrado**

Tipos de hash

Abierto

“Chaining” o “encadenamiento”

Clave	Hash	Valor
S	2	0
E	0	1
A	0	2
R	3	3
C	3	4



Tipos de hash

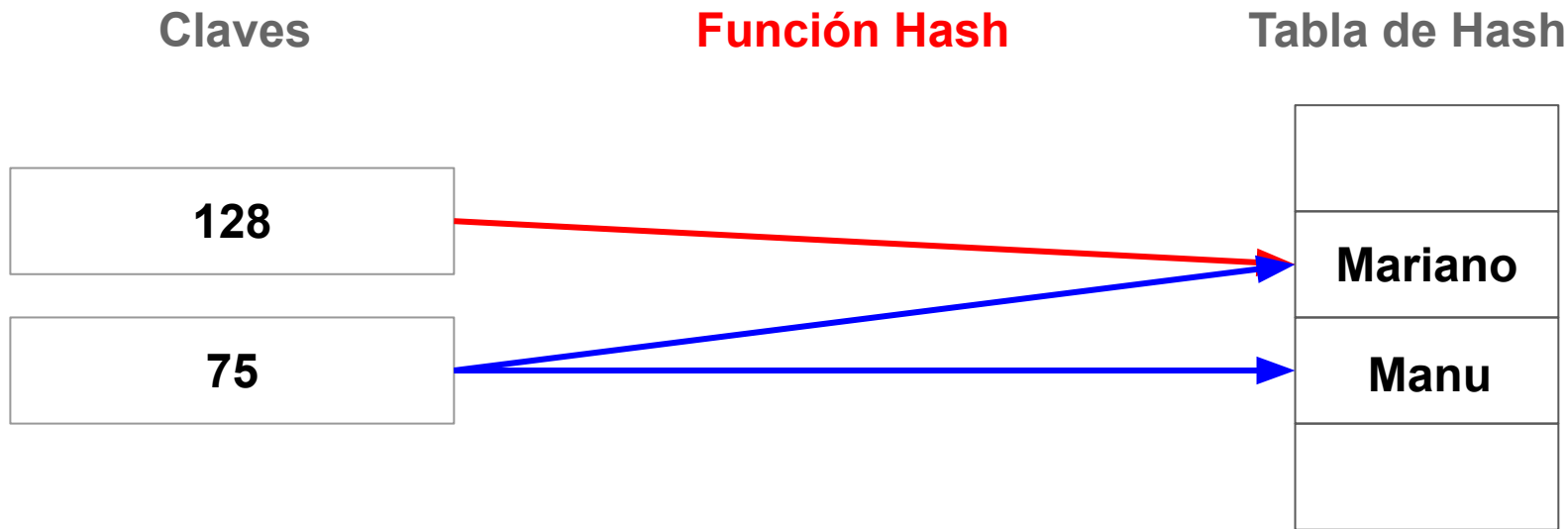
Abierto

- Para encontrar una clave con su correspondiente valor en caso de colisión:
 - $O(n)$
donde n es la cantidad de elementos que colisionaron
 - Es decir, voy a tener que recorrer la lista enlazada

Tipos de hash

Cerrado

- Todos los valores se guardan dentro de la misma tabla
- Tamaño de tabla \geq nro. de claves



Tipos de hash

Cerrado

- Si hay colisión, sigo recorriendo el array hasta encontrar el próximo espacio libre (ver siguiente slide)
- Por esto se conoce como “direccionamiento abierto”

Tipos de hash

Cerrado: Tipos de métodos de búsqueda

- “Probing lineal”: buscar el siguiente espacio libre inmediato
- “Probing cuadrático”: $(\text{intentos fallidos})^2$ para intentar insertar
- Hash doble: aplicar una segunda función de hash a la clave cuando hay colisión

Factor de carga

$$\alpha = n / m$$

n: número de claves almacenadas actualmente

m: capacidad de la tabla de hash

- Indica grado de ocupación de la tabla de hash
 - Y qué tan probable es que haya colisión
- Entre 0 y 1

Rehash

- Cuando $\alpha \geq 0.75$, es hora de rehashear
- La capacidad de la tabla de hash debería aumentar

Ejemplo:

$m = 20$ (capacidad)

$$\begin{aligned}\alpha &= n / m \\ 0.75 &= n / 20 \\ n &= 15\end{aligned}$$

Cuando almacene la clave 15, voy a tener que aumentar la capacidad de la tabla de hash (podría duplicarla, por ejemplo)

Operaciones

- Crear
- Insertar
- Obtener
- Cantidad
- Destruir
- Quitar
- Contiene

Ejemplos - Insertar




Clave	Hash	Valor
54	4	000

“Hash” = Clave % Tamaño_tabla

Tamaño_tabla = 10

0		
1		
2		
3		
4		
5		
6		
7		
8		
9		

Ejemplos - Insertar




Clave	Hash	Valor
54	4	000

“Hash” = Clave % Tamaño_tabla

Tamaño_tabla = 10

0		
1		
2		
3		
4	000	54
5		
6		
7		
8		
9		

Ejemplos - Insertar

Clave	Hash	Valor
54	4	000
 26	6	111

“Hash” = Clave % Tamaño_tabla

Tamaño_tabla = 10

0		
1		
2		
3		
4	000	54
5		
6		
7		
8		
9		

Ejemplos - Insertar

Clave	Hash	Valor
54	4	000
26	6	111

0		
1		
2		
3		
4	000	54
5		
6	111	26
7		
8		
9		

“Hash” = Clave % Tamaño_tabla

Tamaño_tabla = 10

Ejemplos - Insertar

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444

“Hash” = Clave % Tamaño_tabla

Tamaño_tabla = 10

0	222	70
1	333	31
2		
3	444	93
4	000	54
5		
6	111	26
7		
8		
9		

Ejemplos - Insertar con colisión

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555

Método lineal

$$P = (P + 1) \% \text{Tamaño_tabla}$$

$$P = (0 + 1) \% 10 = 1$$

0	222	70
1	333	31
2		
3	444	93
4	000	54
5		
6	111	26
7		
8		
9		

Ejemplos - Insertar con colisión

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555


$$P = (P + 1) \% \text{Tamaño_tabla}$$

$$P = (1 + 1) \% 10 = 2$$

0	222	70
1	333	31
2		
3	444	93
4	000	54
5		
6	111	26
7		
8		
9		23

Ejemplos - Insertar con colisión

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	111	26
7		
8		
9		24

Ejemplos - Insertar existente



Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	111	26
7		
8		
9		25

Ejemplos - Insertar existente

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		26

Ejemplos - Rehash

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666
7	7	777
29	9	888

¿Cuál es el
factor de carga si
inserto la nueva
clave?

$$\alpha = n / m$$

n: 8
m: 10

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7	777	7
8		
9		27

Ejemplos - Rehash

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666
7	7	777
29	9	888


$$\alpha = n / m$$
$$\alpha = 8 / 10$$
$$\alpha = 0,8$$

$$\alpha \geq 0,75$$

¡Aumento el tamaño de la tabla de hash!

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7	777	7
8		
9	888	29

Ejemplos - Obtener



Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

$O(1)$!!!



0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		29

Ejemplos - Obtener

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		30

Ejemplos - Obtener

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

No coincide la clave, no es el valor que busco

Sigo buscando hasta que:

- Lo encuentre
- Encuentre espacio vacío

$$P = (P + 1) \% \text{Tamaño_tabla}$$

$$P = (0 + 1) \% 10 = 1$$

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		31

Ejemplos - Obtener

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

No coincide la clave, no es el valor que busco

Sigo buscando hasta que:

- Lo encuentre
- Encuentre espacio vacío

$$P = (P + 1) \% \text{Tamaño_tabla}$$

$$P = (1 + 1) \% 10 = 2$$


0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		32

Ejemplos - Obtener

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		33

Ejemplos - Obtener inexistente

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666
 3	3	777


No coincide la clave, no es el valor que busco

Sigo buscando hasta que:

- Lo encuentre
- Encuentre espacio vacío

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		34

Ejemplos - Obtener inexistente

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666
 3	3	777


No coincide la clave, no es el valor que busco

Sigo buscando hasta que:

- Lo encuentre
- Encuentre espacio vacío

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		35

Ejemplos - Obtener inexistente

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666
 3	3	777

El espacio se
encuentra vacío

La clave no puede
estar en la tabla,
porque se hubiese
posicionado en ese
espacio

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		36

Ejemplos - Quitar



Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		37

Ejemplos - Quitar

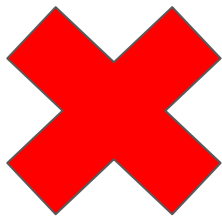
Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		38

Ejemplos - Quitar

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		39



¿Por qué esto estaría mal?

Ejemplos - Quitar

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		40

Ejemplos - Quitar

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

No coincide la clave, no es el valor que busco

Sigo buscando hasta que:

- Lo encuentre
- Encuentre espacio vacío

$$P = (P + 1) \% \text{Tamaño_tabla}$$

$$P = (0 + 1) \% 10 = 1$$

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		41

Ejemplos - Quitar

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

El espacio se
encuentra vacío

La clave no puede
estar en la tabla,
porque se hubiese
posicionado en ese
espacio

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		42

Ejemplos - Quitar

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

Pero...
La clave sí está en
la tabla de hash

...

¿Qué pasó?

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		43

Ejemplos - Quitar

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

Quitamos la clave
31 con “hash” 1...

Si buscamos la
clave 40 con
“hash” 0 pero que
está almacenada
más adelante en
la tabla de hash,
¡tenemos un
problema!

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		44

¿Cómo quitar? Métodos:

- Reemplazar el espacio que acabamos de vaciar
- Utilizar un flag para indicar que se borró algo

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1	333	31
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		46

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		47

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

Avanzo hasta encontrar:

- el próximo espacio vacío
- una clave que pueda ser movida a ese nuevo espacio que vaciamos

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		48

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

Avanzo hasta encontrar:

- el próximo espacio vacío
- una clave que pueda ser movida a ese nuevo espacio que vaciamos

→ Me encuentro con la clave 40 con “hash” 0

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		49

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

El espacio está
ocupado, sigo
recorriendo

$$P = (P + 1) \% \text{Tamaño_tabla}$$

$$P = (0 + 1) \% 10 = 1$$

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		50

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

El espacio está libre,
puedo guardarlo acá

0	222	70
1		
2	555	40
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		51

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

El espacio está libre,
puedo guardarlo acá

Pero lo saco de
acá...

Volvió a quedar un
espacio libre →
aplico el mismo
procedimiento

0	222	70
1	555	40
2		
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		52

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

Avanzo hasta encontrar:

- el próximo espacio vacío
- una clave que pueda ser movida a ese nuevo espacio que vaciamos

0	222	70
1	555	40
2		
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		53


Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

En este caso, el siguiente no es vacío y la clave 93 con "hash" 3 está bien posicionada

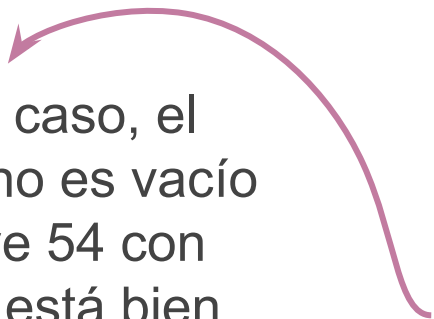
0	222	70
1	555	40
2		
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		54

Ejemplos - Quitar con reemplazo



Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

En este caso, el siguiente no es vacío y la clave 54 con "hash" 4 está bien posicionada

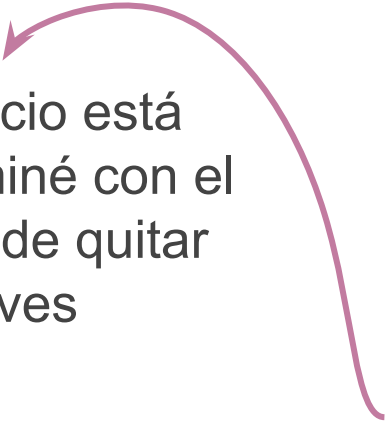


0	222	70
1	555	40
2		
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		55

Ejemplos - Quitar con reemplazo

Clave	Hash	Valor
54	4	000
26	6	111
70	0	222
31	1	333
93	3	444
40	0	555
26	6	666

El espacio está
libre, terminé con el
proceso de quitar
claves



0	222	70
1	555	40
2		
3	444	93
4	000	54
5		
6	666	26
7		
8		
9		56

¿Cómo quitar? Métodos:

- Reemplazar el espacio que acabamos de vaciar
- Utilizar un flag para indicar que se borró algo
 - Sino ocurre lo mismo que vimos antes: si está vacío, asumo que ya no hay nada hasheado más adelante y esto puede traer problemas

¿Preguntas?

