

TABLAS HASH

**Búsqueda mediante transformación de claves
(hashing)**



Tablas Hash

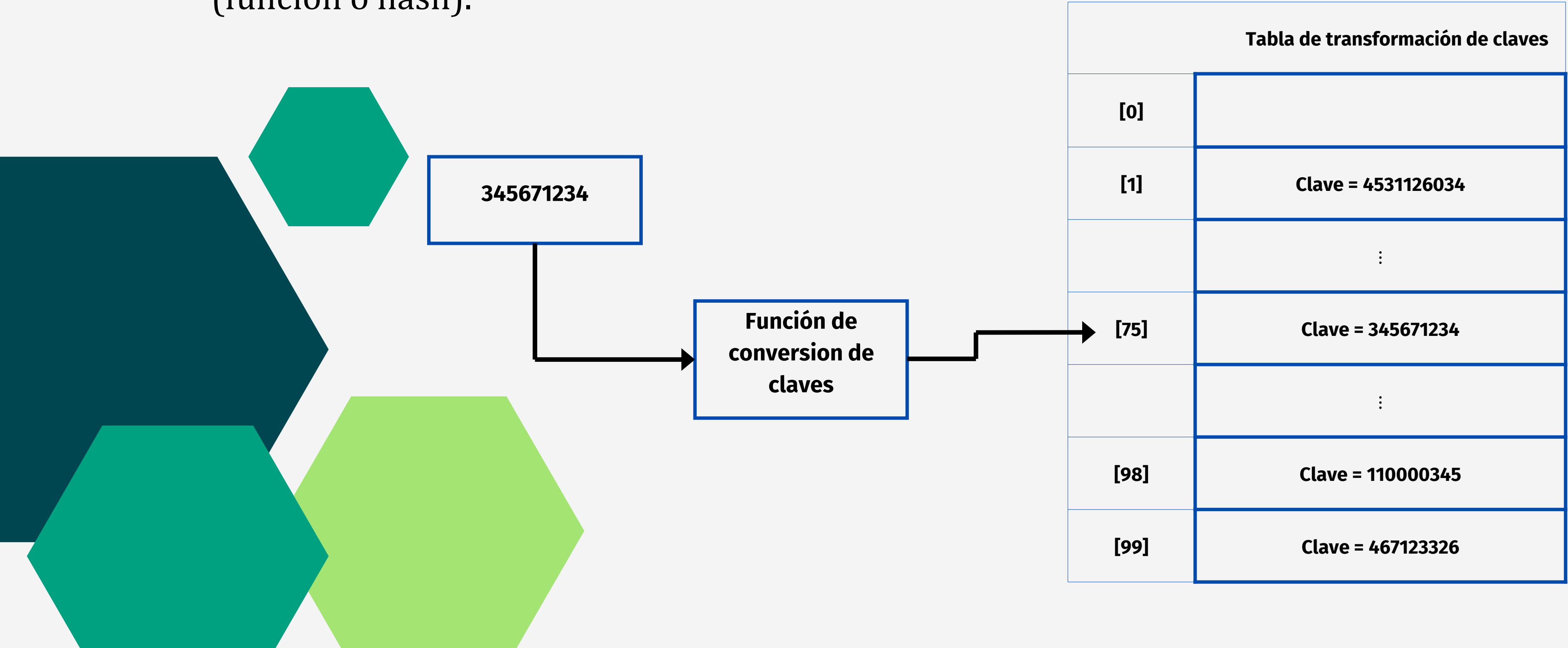
Una **tabla hash**, matriz asociativa, hashing, mapa hash, tabla de dispersión o tabla fragmentada es una estructura de datos que asocia *llaves* o *claves* con *valores*.


La operación principal que soporta de manera eficiente es la *búsqueda*.



El método de transformación de claves consiste en convertir la clave dada (numérica o alfanumérica) en una dirección (índice) dentro del array.

La correspondencia entre las claves y la dirección en el medio de almacenamiento en el array se establece por una función de conversión (función o hash).



- 
- La función de transformación de clave, $H(k)$ convierte la clave (k) en una dirección (d) .
 - La función H es, por consiguiente, una función de paso o conversión de múltiples claves a direcciones.

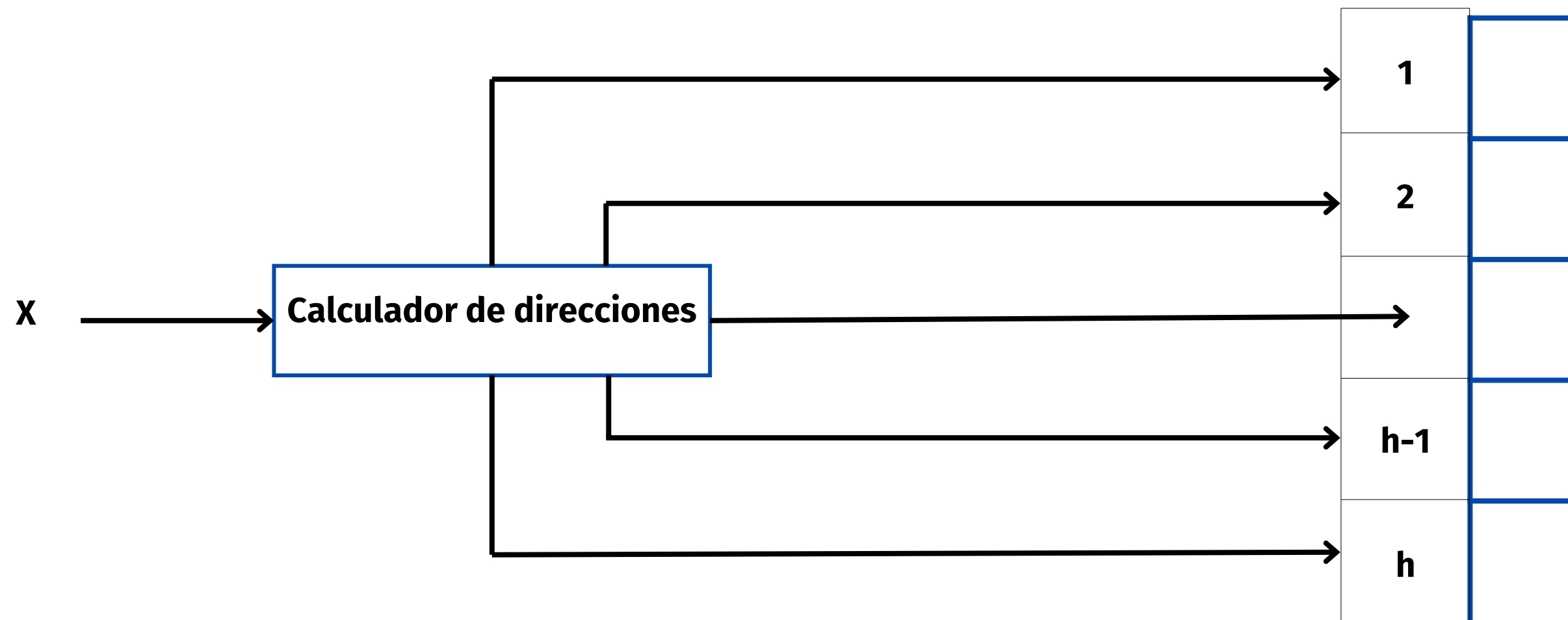
¿Como resolver la situación de que H no produzca la posición del elemento asociado?

- Se debe utilizar algún método para producir una dirección alternativa, es decir, d' , y si esta no es aún la posición del elemento deseado, se produce un tercer índice d'' , y así sucesivamente.
- En el caso en el que una clave distinta de la deseada está en la posición identificada se denomina **colisión**; la tarea de generación de índices alternativos se denomina **tratamiento de colisiones**.
- La elección de una buena función de conversión exige un tratamiento idóneo de colisiones, es decir, la reducción del número de colisiones.

METODO DE TRANSFORMACIÓN DE CLAVES

Los métodos de transformación de claves convierten claves en direcciones. La función de conversión equivale a una caja negra que podríamos llamar **calculador de direcciones**.

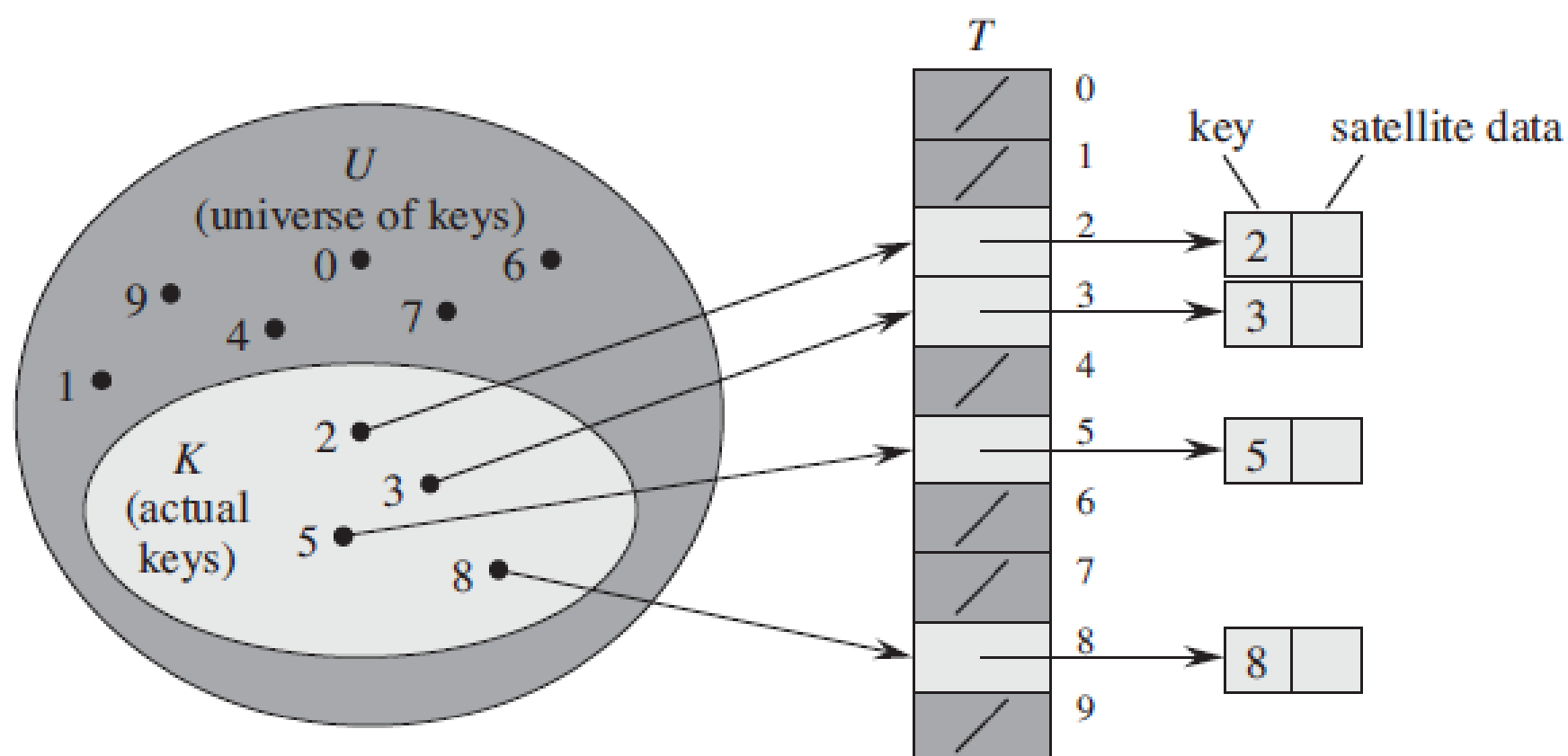
Cuando se desea localizar un elemento de clave X, el indicador de direcciones indicará en que posición del array estará situado el elemento.



Direccionamiento directo

El direccionamiento directo es una técnica simple que funciona bien cuando el universo U de llaves es razonablemente pequeño.

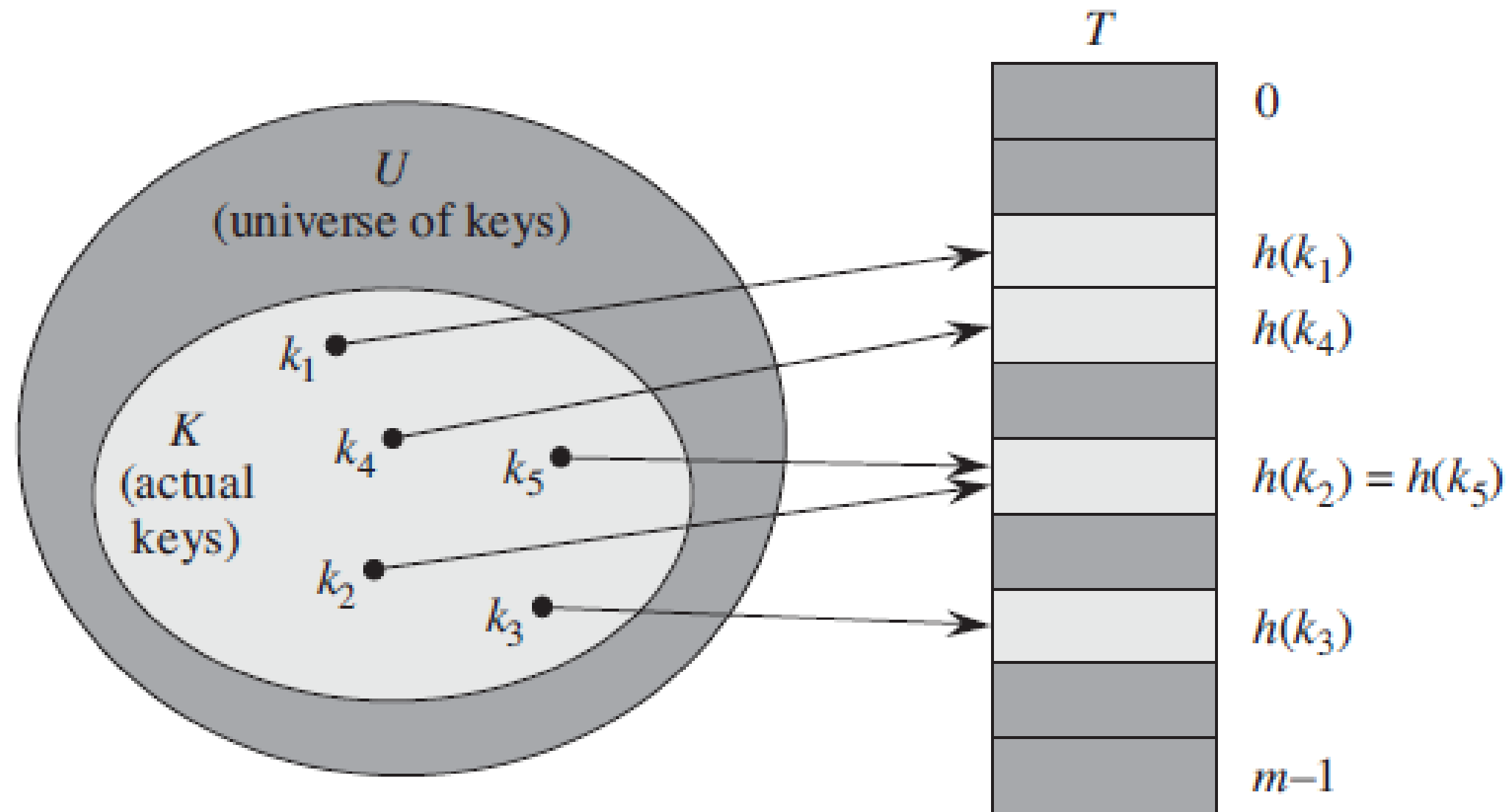
Supondremos que no hay dos elementos con la misma clave.



Cómo implementar un conjunto dinámico mediante una tabla de direcciones directas T .

- ❖ Cada clave del universo $U = \{0, 1, \dots, 9\}$ corresponde a un índice de la tabla.
- ❖ El conjunto $K = \{2, 3, 5, 8\}$ de llaves actuales determina los espacios en la tabla que contienen punteros a elementos.
- ❖ Los otros espacios, más sombreados, contienen *NULO*.

Uso de funciones hash



Uso de una función hash h para asignar claves en índices de la tabla hash.

Note que las claves k_2 y k_5 mapean en el mismo espacio y colisionan.



Resolución de colisiones por encadenamiento.

Cada ranura de tabla hash **T[j]** contiene una lista enlazada de todas las claves cuyo valor hash es **j**.
Por ejemplo, $h(k_1) = h(k_4)$ y $h(k_5) = h(k_7) = h(k_2)$.

La lista enlazada puede estar simple o doblemente enlazada.

Resolución de colisiones por encadenamiento.

Cada ranura de tabla hash **T[j]** contiene una lista enlazada de todas las claves cuyo valor hash es **j**.
Por ejemplo, $h(k_1) = h(k_4)$ y $h(k_5) = h(k_7) = h(k_2)$.

La lista enlazada puede estar simple o doblemente enlazada.

Resolución de colisiones por encadenamiento.

Cada ranura de tabla hash **T[j]** contiene una lista enlazada de todas las claves cuyo valor hash es **j**.
Por ejemplo, $h(k_1) = h(k_4)$ y $h(k_5) = h(k_7) = h(k_2)$.

La lista enlazada puede estar simple o doblemente enlazada.

Funciones Hash



TRUNCAMIENTO

Ignora la parte de la clave y se utiliza la parte restante directamente como índice (considerando campos no numéricos y sus códigos numéricos).

Si las claves son enteros de 8 dígitos y la tabla de transformación tiene 1000 posiciones, entonces el primero, segundo y quinto dígito desde la derecha pueden formar la función de conversión.

Ejemplo:

7 2 5 8⁵ 8 4 9² 5¹ se convierte en 895

El truncamiento es un método muy rápido, pero falla para distribuir las claves de modo uniforme.

PLEGAMIENTO

La técnica de plegamiento consiste en la partición de la clave en diferentes partes y la combinación de las partes en un modo conveniente (a menudo utilizando suma o multiplicación) para obtener el índice.

La clave X se divide en varias partes, X_1, X_2, \dots, X_n , donde cada parte, con la única posible excepción de la última, tiene el mismo número de dígitos que la dirección más alta que podría ser utilizada.

A continuación, se suman todas las partes: $h(X) = X_1 + X_2 + \dots + X_n$

En esta operación se desprecian los dígitos más significativos que se obtengan de arrastre o acarreo.

Ejemplo 1:

Un entero de ocho dígitos se puede dividir en grupos de tres, tres y dos, los grupos se suman juntos y se truncan si es necesario para que estén en el rango adecuado de índices.

Si la clave es:

Clave = 62538194

y el número de direcciones es 100, entonces:

$$625 + 381 + 94 = 1100$$

Que se truncara a 100 y que sera la dirección deseada.



Ejemplo 2:

Los numeros de empleados (claves) constan de cuatro dígitos y las direcciones reales son 100.

Calcular las direcciones correspondientes por el método de plegamiento.

$$h(4205) = 42 + 05 = 47$$

$$h(8148) = 81 + 48 = 129 \text{ y se trunca a } 29$$

$$h(3355) = 33 + 55 = 88$$

ARITMETICA MODULAR

Convertir la clave a un entero, dividir por el tamaño del rango del índice y tomar el resto como resultado. La función de conversión utilizada es mod (modulo o resto de la división entera).

$$h(x) = x \bmod m$$

Donde m es el tamaño del array con índices de 0 a m-1. La mejor elección de los módulos son los números primos.

Por ejemplo, en un array de 1000 elementos se puede elegir 997 o 1009.

Ejemplo 1

$$18 \bmod 6 = 0$$

$$19 \bmod 6 = 1$$

$$20 \bmod 6 = 2$$

0	18
1	19
2	20
3	
4	
5	

m = 6

Si se desea que las direcciones vayan de 0 hasta m, la función de conversión debe ser:

$$h(x) = x \bmod (m+1)$$

Ejemplo 2: La clave de búsqueda es una cadena de caracteres - (como el nombre).

Primero hay que obtener las direcciones de conversion.

El método más simple es asignar a cada carácter de la cadena un valor entero

(por ejemplo; A = 1, B = 2, C =3, ...) y sumar los valores de los caracteres en la cadena.

Si el nombre fue JONAS, esta clave se convertiría en el entero

J					O					N					A					S																													
10					+					15					+					14					+					1					+					19					= 59				

MITAD DEL CUADRADO

Este método consiste en calcular el cuadrado de la clave x . La función se define como $h(x) = c$ donde c se obtiene eliminando dígitos a ambos extremos de x^2 . Se debe utilizar las mismas posiciones de x^2 para todas las claves.

Ejemplo: Una empresa tiene ochenta empleados y cada uno tiene claves de cuatro dígitos y el conjunto de direcciones de memoria va de 0 a 100.

x	4205	7148	3350
x^2	17682025	51093904	11222500
$h(x)$	82	93	22

Si elegimos, por ejemplo, el cuarto y el quinto dígito significativo quedaría como:

