

Algoritmo de Búsqueda Indexada

Los Walo

19 de julio del 2013

Índice

- 1 Introducción
- 2 Complejidad
- 3 Como funciona
- 4 Cuando Ocupamos Búsqueda Indexada
- 5 Ventajas y desventajas
- 6 Conclusiones

Que es y en que consiste

- Algoritmo de Búsqueda Recursivo.
- Uso de Key o Índices.
- Acción de registrar ordenadamente la información para elaborar su índice.
- La indexación es un elemento fundamental para motores de búsqueda y las base de datos.
- Es la manera mas eficiente de organizar y recuperar contenido de una base de datos.

Complejidad

La complejidad es variable puede depender de:

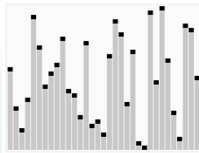
- Algoritmo de Ordenamiento.
- Uso de otros algoritmos de búsqueda.
- Forma de llenado de arreglos de indices.
- Otros.

Calculo de la complejidad notación $O(n)$:

$$O(n) = 7n^2 + 8n + n/c + 1 = n^2 + \log(n) = n^2$$

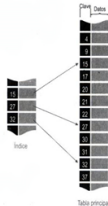
Complejidad (Peor caso $O(n)$)

Ordenamiento
 Quicksort



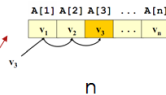
n^2

Algoritmo Bus.
 Indexada



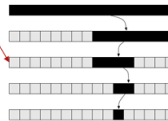
n^2

Secuencial



n

Binaria



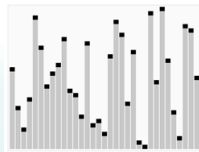
$\log_2(n)$

Final

1. $n^2 + n^2 + n = n^2$
2. $n^2 + n^2 + \log_2(n) = n^2$

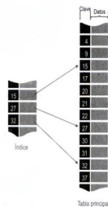
Complejidad (Caso Medio $O(n)$)

Ordenamiento
Quicksort



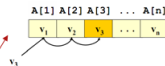
$$n \log_2(n)$$

Algoritmo Bus.
Indexada



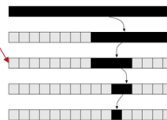
$$n^2$$

Secuencial



$$n/2 = n$$

Binaria



$$\log_2(n)$$

Final

1. $n \log(n) + n^2 + n$
 $= n^2$
2. $n \log 2n + n^2 + \log 2n$
 $= n^2$

Como funciona

Mediante cada elemento del array índice se asocian grupos de elementos del array inicial.

Los elementos en el índice y en el array deben estar ordenados. El método consta de dos pasos:

- Buscar en el array índice el intervalo correspondiente al elemento buscado.
- Restringir la Búsqueda a los elementos del intervalo localizado previamente.

Se puede implementar la búsqueda binaria o secuencial en el array de índices y en el inicial.

Finaliza la búsqueda según las condiciones del algoritmo de búsqueda sub-utilizado (Binario/Secuencial).

Ejemplos (Búsqueda Secuencial)

3	4	5	6	7	8	9	10	14	16	19
---	---	---	---	---	---	---	----	----	----	----

1. Si $X=10$, Buscara en el array de índice comparando la 'clave' con 'X', si 'clave' es ' \geq ' ingresa a buscar al arreglo original de lo contrario continua a la siguiente casilla (clave).
2. Si la 'clave' es ' \geq ' a 'X' ingresara a buscar al sub-arreglo desde la casilla 'posición', de lo contrario continúa comparando en la siguiente casilla (clave).
3. En el array índice la casilla 2 (clave 9) indicara buscar a partir de la casilla 6 del arreglo original hasta el final del bloque.
4. Al ingresar al sub-arreglo [9; 10; 14], ira comparando la igualdad desde la primera casilla a la última del sub-arreglo una tras otra.
5. Al encontrar 10 lo retorna

Clave	→	3	6	9	16
Posición	→	0	3	6	9

Ejemplos (Búsqueda Binaria)

3	4	5	6	7	8	9	10	14	16	19
---	---	---	---	---	---	---	----	----	----	----

1. Si $X=8$, Realiza el mismo tipo de búsqueda en el arreglo de índices que el caso anterior.
2. En el array índice la casilla 1 (clave 6) indicara buscar a partir de la casilla 3 del arreglo original hasta el final del bloque.
3. La búsqueda en el array inicial comenzara en la casilla 3, en el bloque $[6; 7; 8]$
4. Compara si es igual a 'X'. Sino ir a la mitad de arreglo $((n+1)/2)$ y comparara nuevamente, si es mayor seguirá hacia la mitad superior sino hacia la mitad inferior.
5. Al encontrar 8 lo retorna.

Clave	3	6	9	16
Posición	0	3	6	9

Tabla de medición

A[]	1.000	10.000	100.000
Tiempo(s)	0.296 seg.	0,643 seg.	6.079 seg.

Cuando ocupamos Búsqueda Indexada

- Conveniente para archivos con mediana volatilidad, actividad variable y tamaño estable.
- Para N muy grandes (porque?)
- En lugares donde se presente el ingreso de datos (registros) sin ningún tipo de orden específico
- Ejemplos: Spip (que es Spip?)

Ejemplo de Indexación para textos

Texto 1 1 2 3 4 5 6 7 8
El ignorante afirma, el sabio duda y reflexiona

Texto 2 1 2 3 4 5 6 7 8
El sabio no dice todo lo que piensa...

afirma	→	1 [3]
dice	→	2 [4]
duda	→	1 [6]
el	→	1 [1,4], 2 [1]
ignorante	→	1 [2]
lo	→	2 [6]
no	→	2 [3]
piensa	→	2 [8]
que	→	2 [7]
reflexiona	→	1 [8]
sabio	→	1 [5], 2 [2]
todo	→	2 [5]
y	→	1 [7]

Vocabulario

Ocurrencias

Ventajas

- Procesar archivo secuencialmente por orden lógico o al azar.
- Se hace una búsqueda en una tabla de índices pequeña y luego en una parte reducida de la tabla original de registros.

Desventajas

- Implica un aumento en la cantidad de espacio requerido.
- El uso de una lista de índices da una gran sobrecarga de espacio y tiempo para los apuntadores usados en búsquedas de registros.
- La inserción en una tabla secuencial indexada es difícil.
- Los registros deben ser de longitud fija.

Conclusiones

- Es el método de Búsqueda mas rápido, sin embargo necesita ocupar otro tipo de búsqueda.
- Sera más rápida cuando los n sean muy grandes.
- Es uno de los métodos de búsqueda más usado por los motores de búsqueda de Internet.
- Su eficacia se observa en la búsqueda en archivos de gran magnitud.