

Algoritmos externos de ordenamiento

Ordenamiento externo es un término genérico para los algoritmos de ordenamiento que pueden manejar grandes cantidades de información. El ordenamiento externo se requiere cuando la información que se tiene que ordenar no cabe en la memoria principal de una computadora (típicamente la RAM) y un tipo de memoria más lenta (típicamente un disco duro) tiene que utilizarse en el proceso.

Intercalación

En este método de ordenamiento existen dos archivos con llaves ordenadas, los cuales se mezclan para formar un solo archivo. La longitud de los archivos puede ser diferente. El proceso consiste en leer un registro de cada archivo y compararlos, el menor es almacenando en el archivo de resultado y el otro se compara con el siguiente elemento del archivo si existe. El proceso se repite hasta que alguno de los archivos quede vacío y los elementos del otro archivo se almacenan directamente en el archivo resultado.



Mezcla directa

Algoritmo de Mezcla Directa:

Dividir una secuencia inicial de datos en dos sub-cadenas y mezclar elemento a.

Elemento de forma ordenada.

El proceso se repite hasta que la secuencia inicial queda totalmente ordenada.

Pasos:

1. Se divide la secuencia inicial de datos del fichero a en dos mitades b y c
2. Se mezclan b y c combinando elementos aislados para formar pares ordenados
3. La secuencia resultante se almacena en el fichero a y se repiten los pasos 1 y 2 para formar cuádruplos ordenados.
4. Se repiten los pasos anteriores para formar octetos ordenados, y así sucesivamente.
5. Este método de ordenamiento realiza sucesivamente una partición y una fusión que produce secuencias ordenadas de longitud cada vez mayor.
6. En la primera pasada la partición es de 1 y la fusión produce secuencias de longitud 2. Las particiones y las fusiones doblan su tamaño en cada pasada del procesamiento hasta lograr una partición del tamaño del archivo.

09	75	14	68	29	17	31	25	04	05	13	18	72	46	61
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Partición

09	14	29	31	04	13	72	61
----	----	----	----	----	----	----	----

75	68	17	25	05	18	46
----	----	----	----	----	----	----

Fusión

09	75	14	68	17	29	25	31	04	05	13	18	46	72	61
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Partición

09	75	17	29	04	05	46	72
----	----	----	----	----	----	----	----

14	68	25	31	13	18	61
----	----	----	----	----	----	----

Mezcla natural

Es una mejora del algoritmo de mezcla directa puesto que en vez de considerar tramos de tamaño fijo se toman en cuenta para la ordenación en todo momento tramos de longitud máxima. Al igual que la mezcla directa se debe hacer un proceso de partir el archivo original para mezclarlo, posteriormente mientras en el archivo C haya elementos a mezclar.

La idea central de este algoritmo consiste en realizar las particiones tomando secuencias ordenadas de máxima longitud en lugar de secuencias de tamaño fijo previamente determinadas. Luego se realiza la fusión de las secuencias ordenadas, en alternada, sobre dos archivos. Aplicando estas acciones en forma repetida se logrará el archivo original quede ordenado. Para la realización de este proceso de ordenación se necesitarán cuatro archivos. El archivo original F y tres archivos auxiliares a los que se denominara F1, F2 y F3. De estos archivos, dos serán considerados de entrada y dos de salida; esto, de manera alternada, con el objeto de realizar la fusión-partición. El proceso termina cuando en la realización de una fusión-partición el segundo archivo quede vacío. Supongamos que se desea ordenar las claves del archivo F utilizando el método mezcla equilibrada.

F	09	75	14	68	29	17	31	25	04	05	13	18	72	46	61
---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Partición inicial

F2	09	75	29	25	46	61
----	----	----	----	----	----	----

F3	14	68	17	31	04	05	13	18	72
----	----	----	----	----	----	----	----	----	----

Primera fusión-partición

F	09	14	68	75	04	05	13	18	25	46	61	72
---	----	----	----	----	----	----	----	----	----	----	----	----

F1	17	29	31
----	----	----	----

Segunda fusión-partición

F	09	14	17	29	31	68	75
---	----	----	----	----	----	----	----

F1	04	05	13	18	25	46	61	72
----	----	----	----	----	----	----	----	----

Tercera fusión-partición

F2	04	05	09	13	14	17	18	25	29	31	46	61	68	72	75
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

F3 Archivo vacío

F archivo original

F1, F2 y F3 archivos auxiliares para los procesos de partición y fusión

Inicio {

 particionInicial (F, F2, F3) // particiona el archivo original en F2 y F3

 repetir

 alternar procesos{

 particionFusion(F2, F3, F, F1) // particion-fusion de F2,F3 en F y F1

 particionFusion(F, F1, F2, F3) // particion-fusion de F,F1 en F2 y F3

 }

 hasta (eof(F1) || eof(F3))

}

Referencias

1. Fernando A. Lagos B. (2007). Algoritmos de ordenamiento. México: Copyleft.
2. Ricardo Ferris. (1996). Algoritmos y estructuras de datos 1. México: McGraw-Hill.
3. Robert Sedgewick, "Algorithms in C", (third edition), Addison-Wesley, 2001
4. Thomas Cormen et al, "Algorithms", (eighth edition), MIT Press, 1992.