

## Apareo de archivos

El apareo es un proceso en el cual dos archivos que tienen un elemento en común ( se lo denomina clave) y que además están ordenados de la misma forma por dicho elemento, son comparados componente a componente con distintos propósitos.

Supongamos que para realizar el apareo utilizamos dos archivos de registros que están ordenados en forma ascendente, entonces el proceso consiste en leer un registro de cada archivo, comparar el campo que tienen en común para seleccionar el registro de menor clave, procesar el registro seleccionado y leer un nuevo registro del archivo del cual provino el registro procesado.

Esta secuencia se repite mientras no se haya llegado al fin en alguno de los dos archivos, cuando esto suceda, dependiendo del problema que se trate, se finalizará el apareo o se continuará leyendo y procesando los registros que hayan quedado.

Consideremos, por ejemplo, este problema:

Una institución educativa tiene dos archivos, uno para cada materia, con la siguiente información de todos los alumnos que rindieron examen final de FÍSICA y de ANÁLISIS : número de legajo, nombre y nota obtenida, ambos archivos están ordenados en forma ascendente por número de legajo. Se solicita que utilizando la información contenida en dichos archivos se emita un listado, ordenado de menor a mayor por número de legajo, con número de legajo y una de estas leyendas:

“Rindio Física” , “Rindio Analisis” o “Rindio Física y Analisis”

Del análisis del problema surge que se pueden presentar tres casos cuando se compare el número de legajo de un alumno que rindió Análisis con el número de legajo de un alumno que rindió Física:

- a) que sea menor
- b) que sea mayor
- c) que sea igual

El caso **a)** sucede cuando el alumno de menor legajo rindió sólo Análisis entonces es su legajo el que debe aparecer en el listado acompañado de la leyenda correspondiente.

El caso **b)** se da cuando el alumno de menor legajo rindió sólo Física entonces es su legajo el que debe aparecer en el listado.

El caso **c)** se verifica si el alumno rindió final de ambas materias entonces se elige uno de los legajos para enviar al listado.

El ordenamiento de los archivos determina cuál de los registros que intervienen en la comparación será procesado primero, si el orden es ascendente la prioridad corresponde al de menor clave, en cambio, si el orden es descendente el que tiene prioridad es el de mayor clave.

En la estrategia de resolución del problema (ver página 2) pueden observarse tres ciclos:



Este ciclo se utiliza mientras hay registros en los dos archivos que se aparean, en cuanto uno de los dos archivos finaliza o si finalizan los dos simultáneamente, se sale del ciclo.



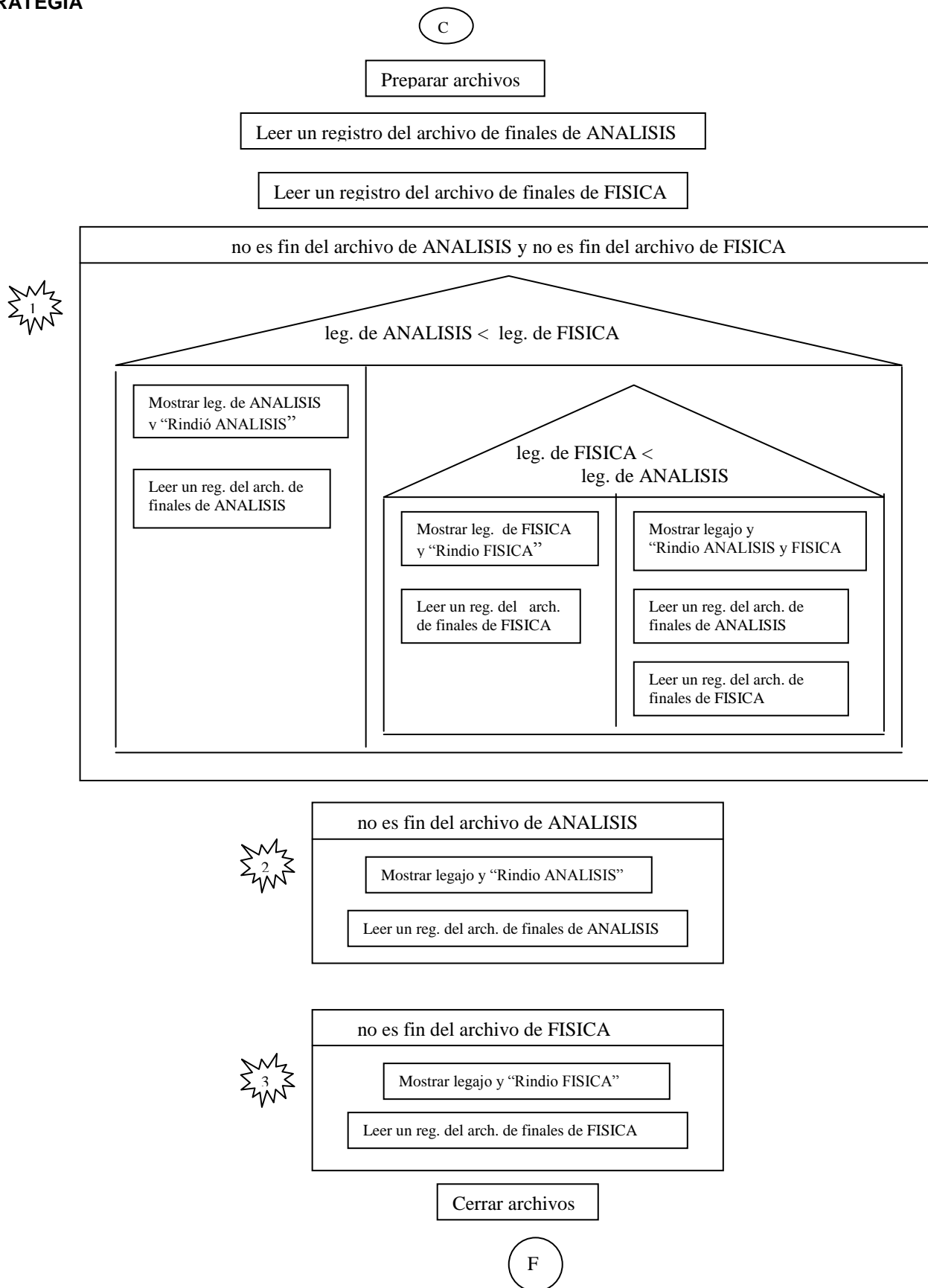
Si la condición de este ciclo es verdadera significa todavía hay registros en el archivo de finales de Análisis y por lo tanto deben procesarse hasta llegar al final.

Si desde un principio la condición del ciclo es falsa entonces éste no se realiza porque ya no hay más registros en el archivo.



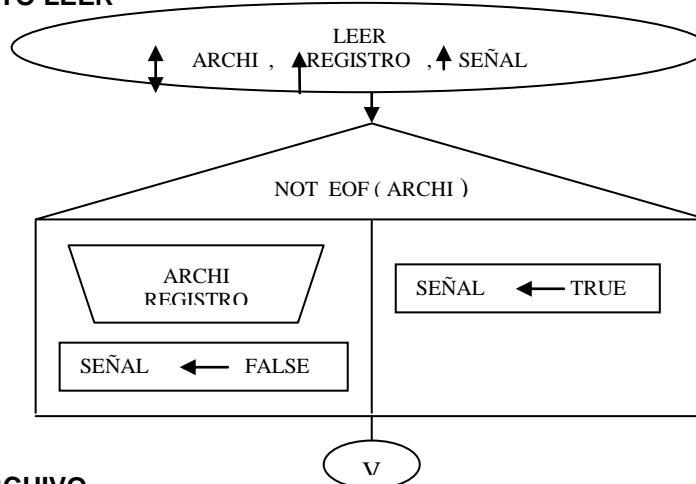
Este ciclo realiza la misma tarea que el ciclo anterior pero para el archivo de finales de Física.

## ESTRATEGIA



**ALGORITMO**

Como podrá observarse en la estrategia, el algoritmo del apareo requiere que el ingreso de datos se haga en distintos lugares del mismo, considerando que para la codificación se empleará el lenguaje Pascal y en este lenguaje es necesario preguntar por el fin de archivo antes de leer, entonces para leer de un archivo se usará el procedimiento LEER que tiene tres parámetros: un archivo, un registro y una señal o bandera. Si no se llegó al fin de archivo, el procedimiento lee un registro y lo devuelve junto con una señal en falso (FALSE). Si se llegó al fin de archivo, no se lee ningún registro pero se devuelve la señal en verdadero(TRUE).

**ALGORITMO DEL PROCEDIMIENTO LEER****DISEÑO DEL REGISTRO DEL ARCHIVO**

TYPE REGALU = RECORD

LEGAJO : longint ; NOMBRE : string [ 20 ]; NOTA : byte

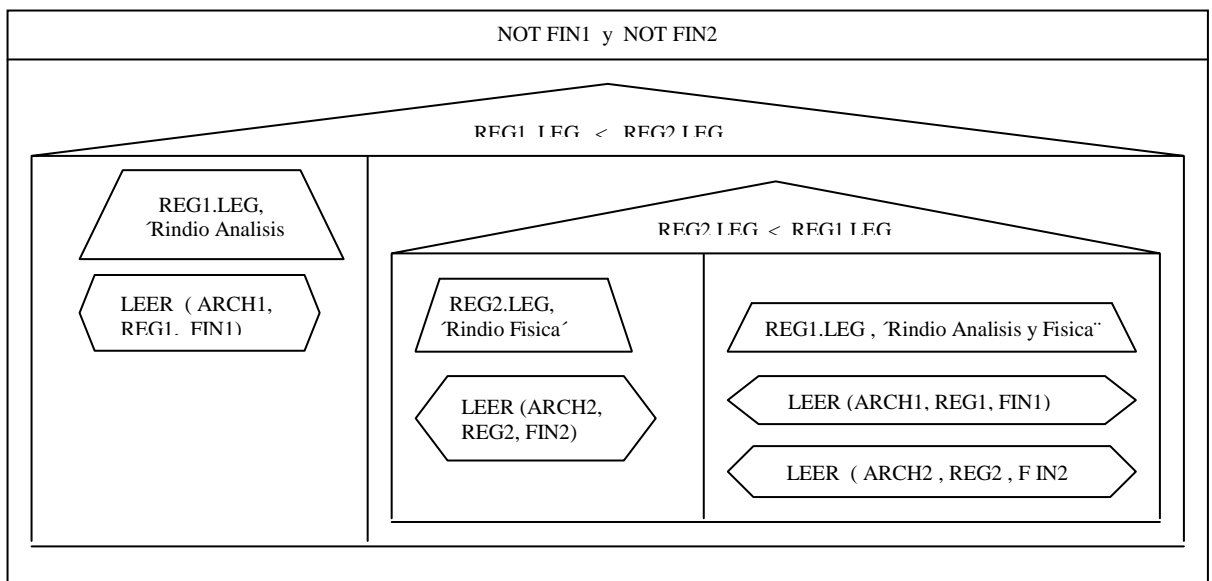
END ;

C

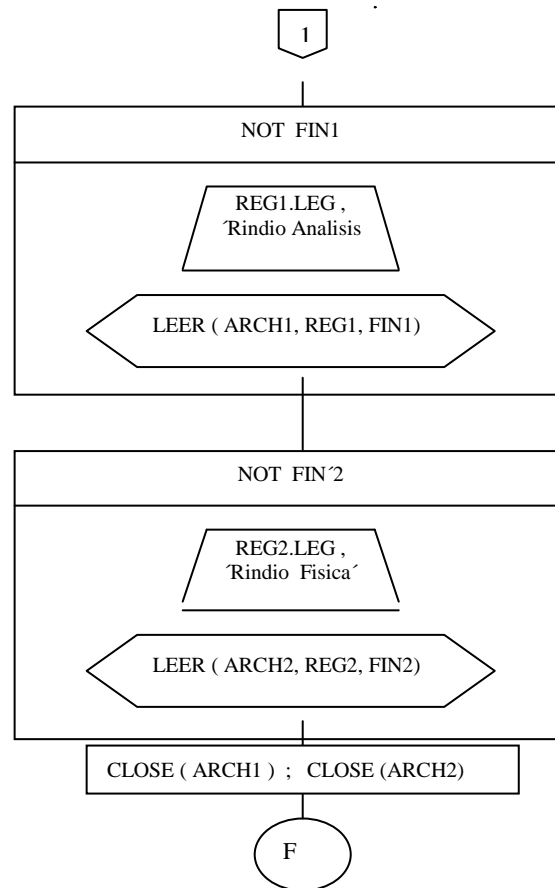
ASSIGN ( ARCH1 , 'ANALISIS.DAT' ) ; RESET ( ARCH1 ) ; ASSIGN(ARCH2, 'FISICA.DAT') ; RESET (ARCH2)

LEER ( ARCH1 , REG1 , FIN1 )

LEER ( ARCH2 , REG2 , FIN2 )



1

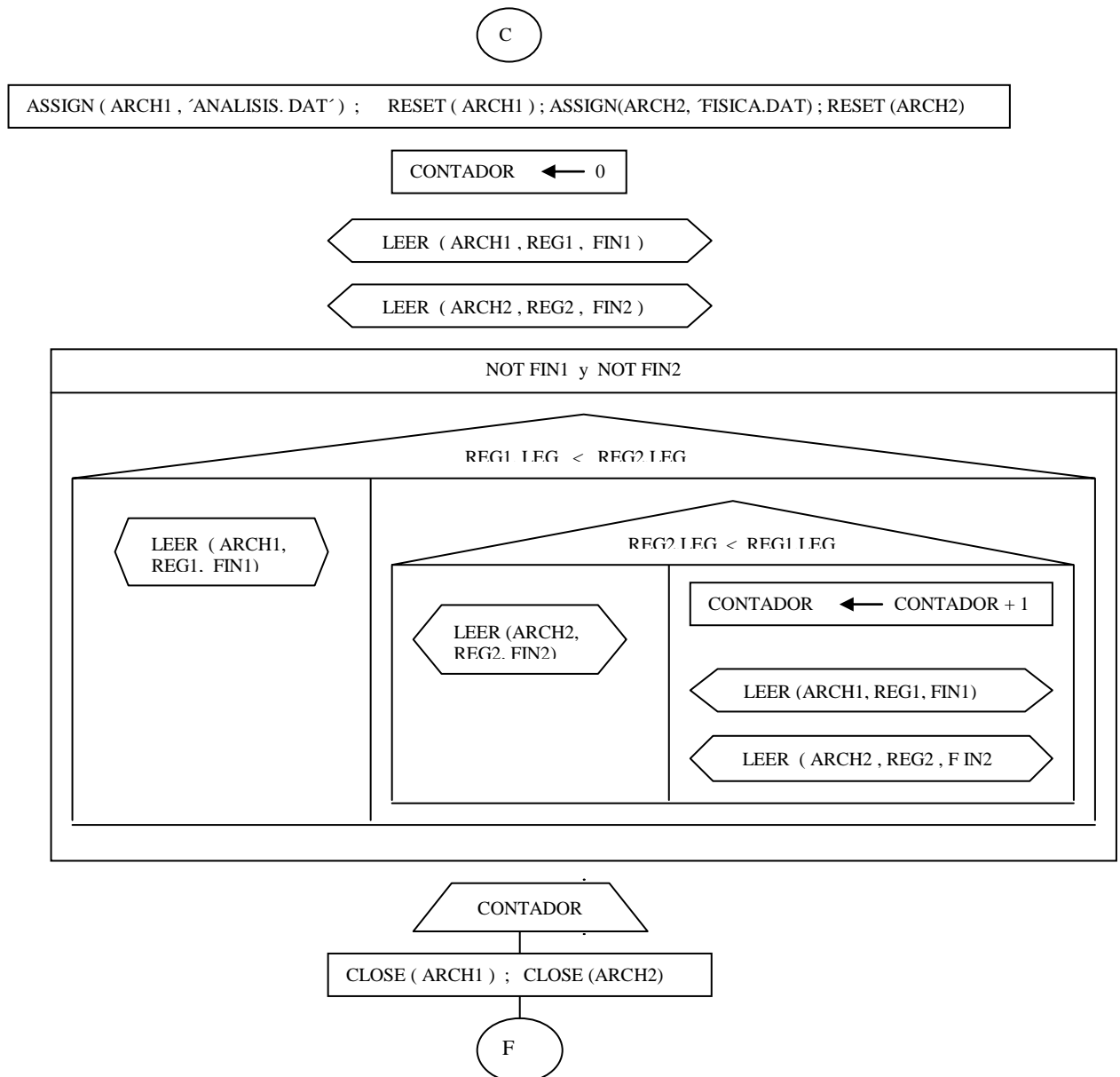


Veamos ahora la utilización del algoritmo en otro problema:

Una institución educativa tiene dos archivos, uno para cada materia, con la siguiente información de todos los alumnos que rindieron examen final de FÍSICA y de ANÁLISIS: número de legajo, nombre y nota obtenida, ambos archivos están ordenados en forma ascendente por número de legajo. Se solicita que utilizando la información contenida en dichos archivos informe cuántos alumnos rindieron los dos finales.

En el caso que se plantea nos interesan los alumnos que estén en ambos archivos, por lo tanto sólo utilizaremos el ciclo 1 del algoritmo porque en él se procesa la información de los dos archivos, cuando se sale de este ciclo si en alguno de los archivos quedaron registros sin procesar no los tenemos en cuenta porque corresponden a alumnos que rindieron un único final.

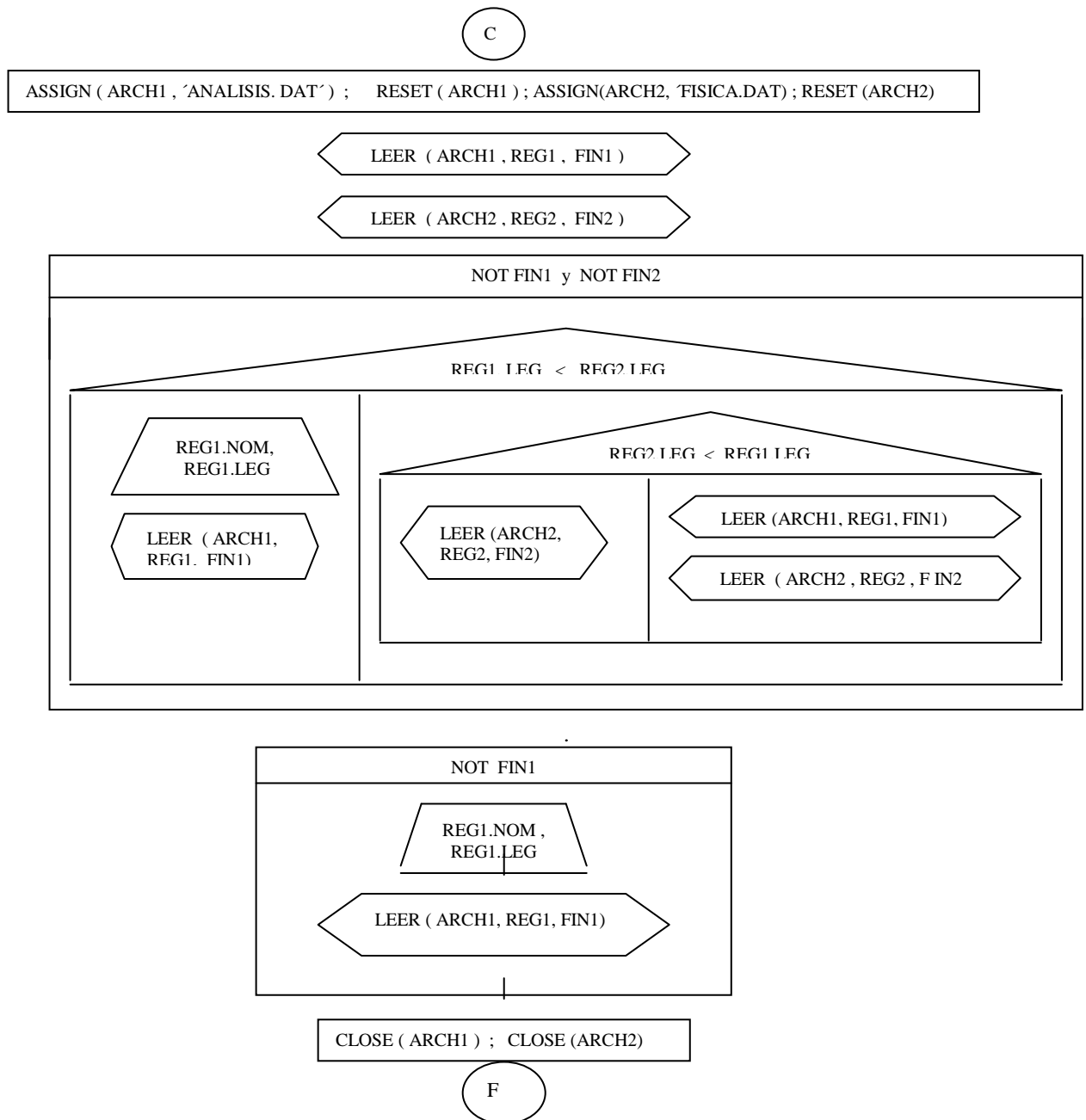
Dentro del ciclo, la comparación de los legajos es necesaria para ir avanzando por los archivos en forma ordenada sin saltar ningún registro, obsérvese que cuando los legajos no coinciden la única acción que se realiza es la de leer del archivo que corresponda.



Para finalizar, nos ocuparemos de un problema que también utiliza parcialmente el algoritmo del apareo:

Una institución educativa tiene dos archivos, uno para cada materia, con la siguiente información de todos los alumnos que rindieron examen final de FÍSICA y de ANÁLISIS : número de legajo, nombre y nota obtenida, ambos archivos están ordenados en forma ascendente por número de legajo. Se solicita que utilizando la información contenida en dichos archivos muestre el nombre y legajo de los alumnos que rindieron Análisis pero no Física.

En este problema utilizaremos el ciclo 1 para detectar a los alumnos que únicamente estén en el archivo de la materia Análisis y el ciclo 2 para listar los registros que pudieran haber quedado en el archivo de Análisis.



## RESUMEN

- El apareo de archivos es un proceso que se realiza con dos archivos que están ordenados por la misma clave y en el mismo sentido.
- Se empieza leyendo un registro de cada archivo y luego, dependiendo del resultado de la comparación de las claves, se continúa leyendo de uno de los archivos o de los dos.
- Según lo que se desee hacer por medio del apareo, el algoritmo podrá tener uno, dos o tres ciclos indefinidos.
- Algunos de los problemas que se pueden resolver por medio de un apareo son: mezcla de archivos, intersección de archivos, actualización de archivos por altas, bajas y modificaciones, etc.
- Si como resultado del apareo se obtienen archivos o listados éstos tienen el mismo orden que los archivos de entrada.