

MÓDULO 6 | ARCHIVOS - ALGORITMOS DE PROCESAMIENTO

1. APAREO

Así como el corte de control nos sirve para generar un reporte, el **apareo** nos sirve para asociar/relacionar datos que se encuentran en **distintos** archivos.

La idea básica es: a partir de dos archivos (uno principal y otro relacionado) que tienen alguna información que los enlace, generar un tercero (o una salida por pantalla), como una mezcla de los dos.

Para hacer esto **es necesario** que ambos archivos estén **ordenados por el valor que los relaciona**.

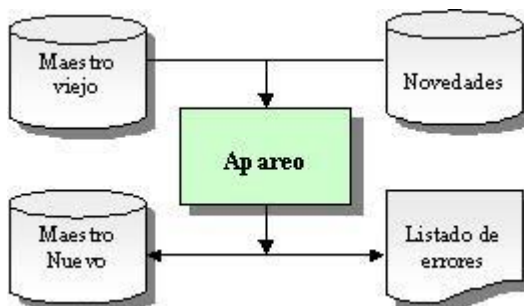
El **APAREO** es una técnica de procesamiento que involucra **dos archivos** que tienen un **elemento en común** (se lo denomina **clave**) y que además están **ordenados** de la misma forma por dicho elemento, son **comparados** componente a componente con distintos propósitos: generar una salida combinada a partir de estos dos archivos.

Este proceso es muy empleado para la **actualización del maestro** a través del archivo de novedades. Las novedades tendrán que ver con **altas, bajas o modificaciones** o algunas de ellas solamente. Si ambos archivos poseen una organización secuencial, entonces ambos archivos deben encontrarse ordenados –ascendente o descendente– por medio del valor de una clave en común.

El resultado de este proceso será un nuevo archivo de salida con la misma estructura que el maestro a actualizar, siendo este archivo el maestro actualizado. El archivo de novedades tiene la misma estructura que el maestro pero con un campo más, el cual indica el *código de movimiento*.

Cada archivo trabajará con sus propios registros, es decir, un registro para el archivo maestro viejo, otro para el archivo de maestro nuevo y un registro para el archivo de novedades. Las situaciones de errores por alta existente o bajas o modificaciones inexistentes, se emitirán por medio del dispositivo de la impresora. Este proceso que genera un nuevo archivo –el maestro actualizado– se denomina **proceso Padre – Hijo**.

El siguiente diagrama de sistema muestra el proceso Padre – Hijo.

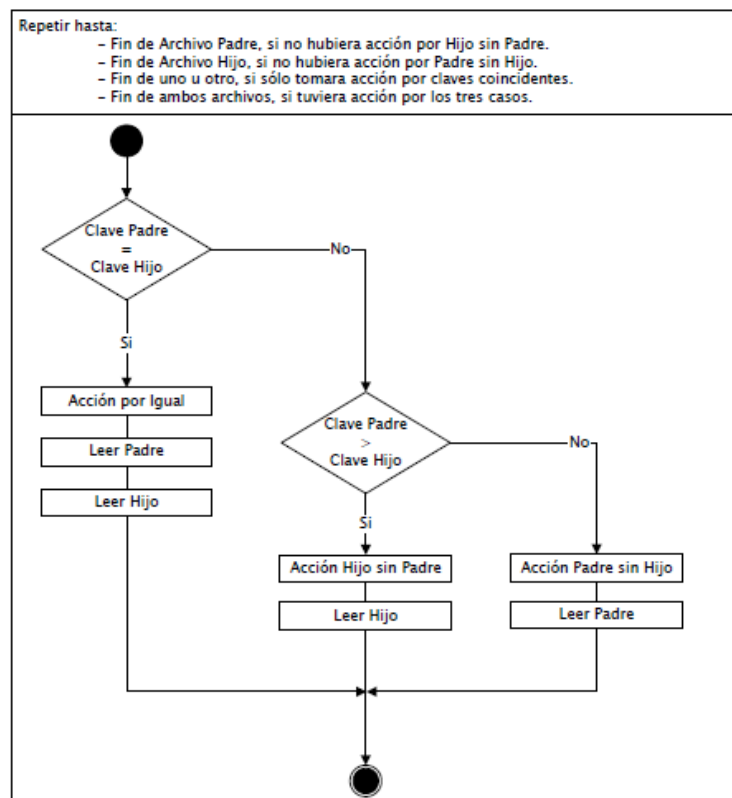


Supongamos que para realizar el apareo utilizamos dos archivos de registros que están **ordenados en forma ascendente**, entonces el proceso consiste en leer un registro de cada archivo, comparar el campo que tienen en común para seleccionar el registro de menor clave, procesar el registro seleccionado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

Dependiendo de la cantidad de archivos que apareemos, clasificaremos el tipo de apareo de la siguiente manera:

1. Uno a uno
2. Uno a muchos
3. Muchos a muchos

APAREO UNO A UNO



Consideremos, por ejemplo, este problema:

Enunciado del 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: "Rindió Física", "Rindió Análisis" o "Rindió Física y Análisis"

Entrada

Archivo Análisis (ordenado secuencial ascendente por NROLEGAJO)

NROLEGAJO	NOMBRE	NOTA
-----------	--------	------

Archivo Física (ordenado secuencial ascendente por NROLEGAJO)

NROLEGAJO	NOMBRE	NOTA
-----------	--------	------

Salida

NRO. LEGAJO	SITUACION
XXX	Rindió Análisis
XXX	Rindió Física
XXX	Rindió Análisis y Física
XXX	Rindió Física
.	.
.	.
.	.
XXX	Rindió Análisis y Física

Análisis del problema

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:

- que sea menor,
- que sea mayor,
- que sea igual

El caso **a)** sucede cuando el alumno de menor legajo rindió sólo Análisis, entonces su legajo es 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.

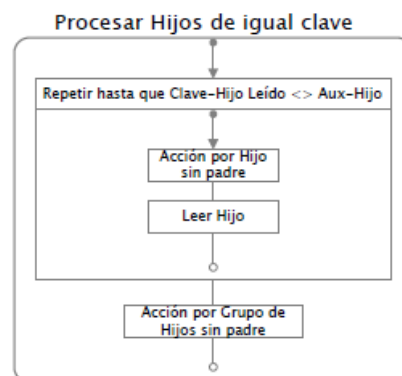
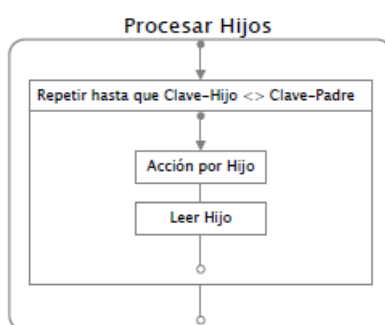
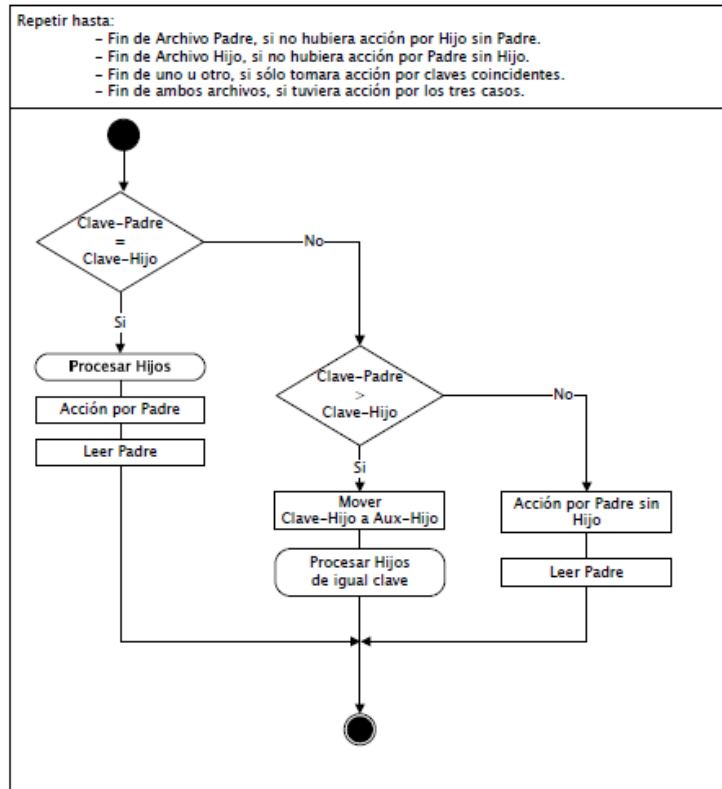
El caso **c)** se verifica si el alumno rindió final de ambas materias entonces se elige uno de los legajos para enviar al listado.

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 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

APAREO UNO A MUCHOS



APAREO MUCHOS A MUCHOS

Repetir hasta:

- Fin de Archivo Entrada1, si no hubiera acción por Entrada2 sin Entrada1.
- Fin de Archivo Entrada2, si no hubiera acción por Entrada1 sin Entrada2.
- Fin de uno u otro, si sólo tomara acción por claves coincidentes.
- Fin de ambos archivos, si tuviera acción por los tres casos.

