



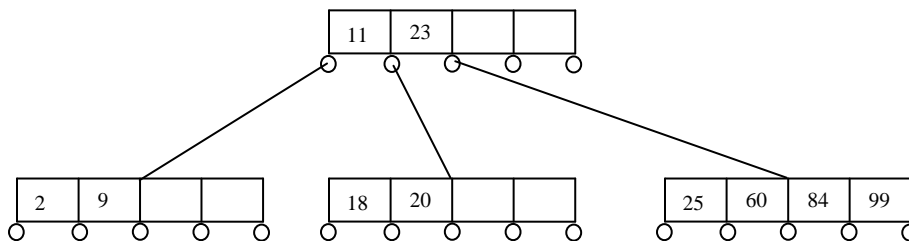
**MATEMÁTICAS PARA LA COMPUTACIÓN**  
**CAPÍTULO 8. ÁRBOLES**

**ÁRBOLES B**  
AUTOR: JOSÉ ALFREDO JIMÉNEZ MURILLO

## ÁRBOLES B

A finales de los años 60 Rodolf. Bayer y Eduard. Mc Creight crearon una nueva forma de organizar información en archivos externos a la cual llamaron árboles B.

La siguiente figura muestra un árbol de orden 4 y profundidad 2.



Las características de los árboles B son:

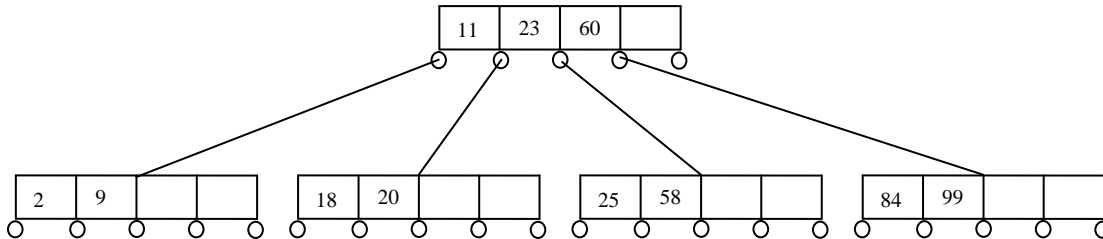
- El orden de un árbol B ( $m$ ) es el número máximo en que se puede partir un nodo. Cada una de estas partes puede contener una clave o dato que se encuentran ordenados en el mismo nodo.
- Si  $m$  es el número de elementos que puede tener un nodo, de cada uno de los nodos pueden partir  $n=(m+1)$  ramas como máximo.
- El árbol está ordenado.
- Todos los nodos terminales (hojas) están en el mismo nivel (Tienen la misma profundidad).
- Todos los nodos intermedios excepto la raíz deben tener mínimo (entero  $(n/2+0.5)$ ) ramas no nulas y máximo  $n$ .
- Todos los nodos intermedios excepto la raíz deben tener como mínimo (entero  $(m/2)$ ) elementos.
- La profundidad ( $h$ ) es el número máximo de consultas para encontrar una clave.

Normalmente se intenta limitar la profundidad del árbol de modo que el número de consultas sea lo más pequeño posible. Mientras más grande sea el orden, menor será la profundidad.

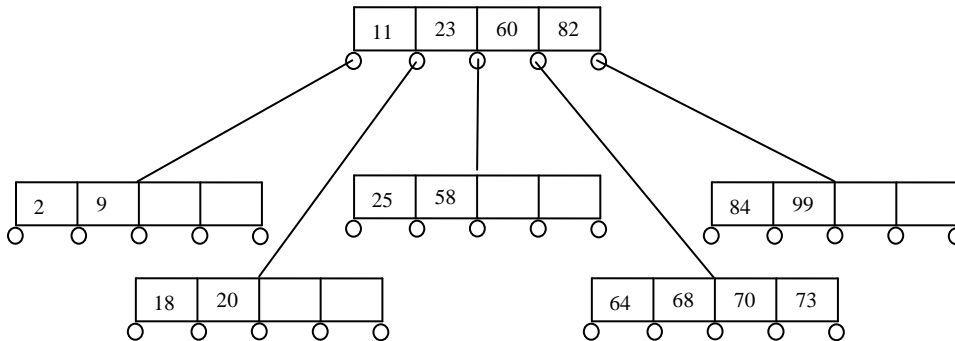
Cuando se crean árboles B para trabajar con archivos de datos en disco, es conveniente tener en cuenta también las características del disco. Para calcular el orden, se utiliza como información fundamental el "cluster". Un cluster es el bloque de disco más pequeño que se lee o se escribe en una misma operación de acceso a disco. Esto permite que el acceso a disco sea más rápido, ya que se evita el desplazamiento mecánico de las cabezas lectoras.

Suponer que se desea insertar el dato 58 en el árbol anterior. De acuerdo a como se están acomodados el dato debería estar en el nodo que tiene los números 25, 60, 84 y 99. Pero no es posible ponerlo ahí porque sus cuatro lugares están llenos. Observar también que si se insertara en ese nodo, el orden de los datos es: 25, 58, 60, 84, 99, en donde el elemento intermedio de estos

cinco datos es el 60, de tal manera que dicho número se debe considerar como padre de los dos nodos en que debe partirse, como se muestra en la siguiente figura.

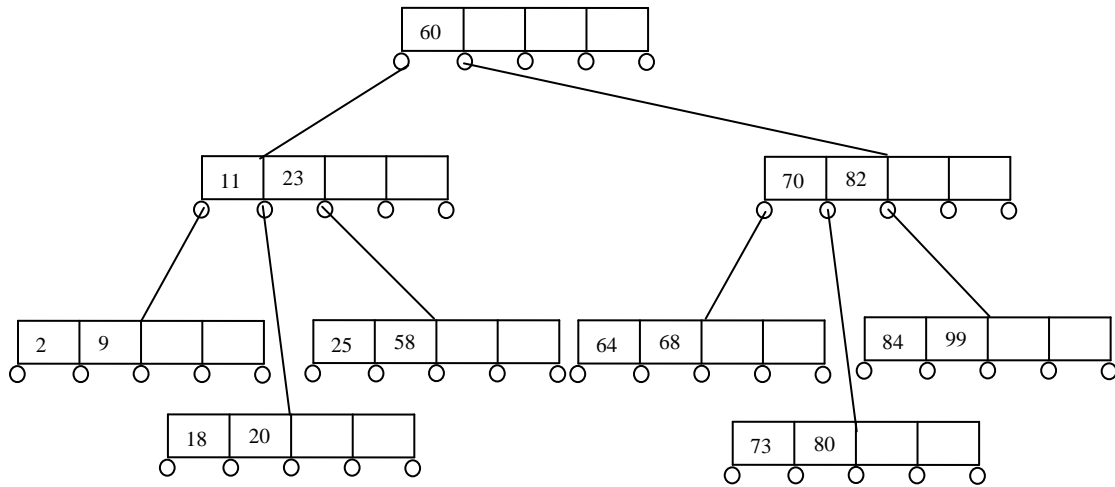


Considerar que se dieron de alta varios datos, de tal forma que el árbol B quedó de la siguiente manera:



Nota: Es importante mencionar que todas las hojas están al mismo nivel ya que es una propiedad de los árboles en B. (Aunque por espacio en este caso se muestran dos nodos un poco más abajo, pero nosotros los debemos considerar con la misma profundidad)

Si se desea insertar el número 80 se puede observar que el lugar que le corresponde es con los datos del nodo que tiene a 64, 68, 70 y 73. De tal manera que el intermedio de 64, 68, 70, 73 y 80 es el 70 que a su vez debe ser el padre de los nuevos dos nodos. Pero el 70 ahora se integra al nodo raíz que tiene los datos 11,23, 60 y 82. Nuevamente no cabe ya que las cuatro localidades del nodo raíz están ocupadas. Por lo tanto se debe partir en dos nuevos nodos y colocar como padre al dato intermedio entre 11, 23, 60, 70, 82 de tal manera que el árbol debe crecer en un nivel como se muestra a continuación.

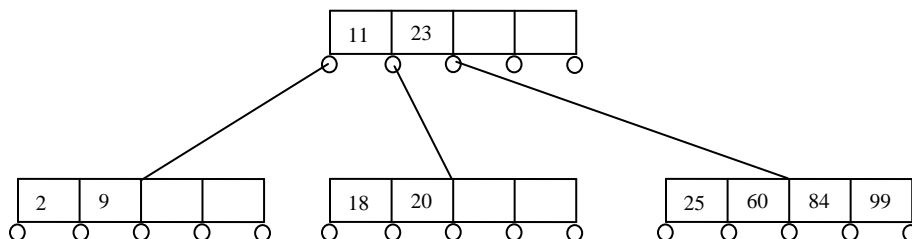


En este árbol de orden 4 que se está tratando, se puede observar que el número:

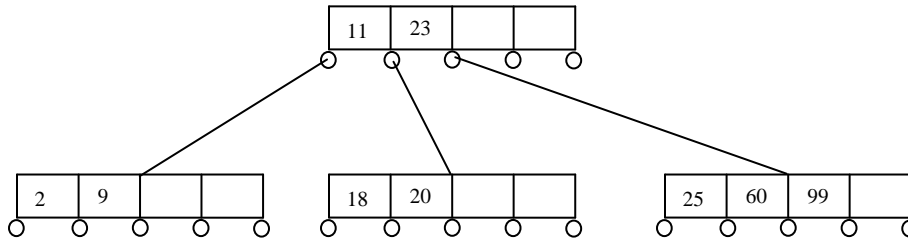
- Máximo de datos por nodo es el orden  $m$  (en este caso  $m = 4$ ).
- Máximo de ramas ( $n$ ) por nodo (en este caso  $n = m+1 = 4+1 = 5$ )
- Mínimo de datos por nodo ( $\text{entero}(m/2)$ ). En este caso ( $\text{entero}(m/2) = \text{entero}(4/2) = 2$ )
- Mínimo de ramas no nulas por nodo ( $\text{entero}(n/2+0.5)$ ). En este caso ( $\text{entero}(n/2+0.5) = \text{entero}(5/2+0.5) = 3$ )

**Borrar información en un árbol en B.**

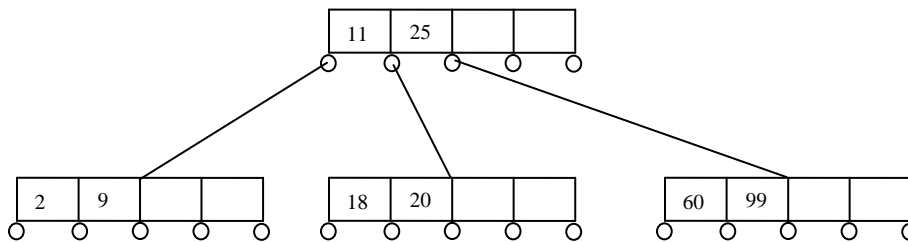
Para borrar un dato que se encuentra en una hoja en donde el número de elementos después de llevar a cabo la baja es mayor que el mínimo requerido solamente se elimina el dato y se recorre la información en caso de ser necesario. Considerar que se tiene el siguiente árbol B.



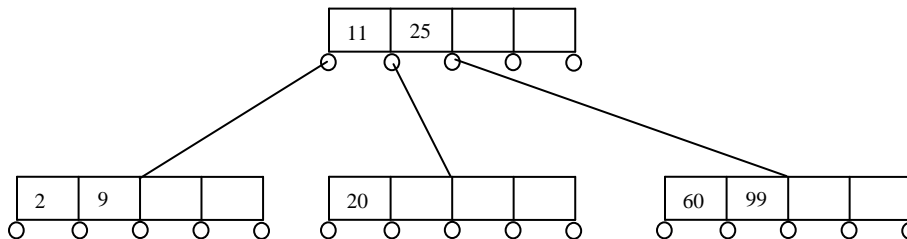
Después de dar de baja el elemento 84 el árbol queda de la siguiente forma.



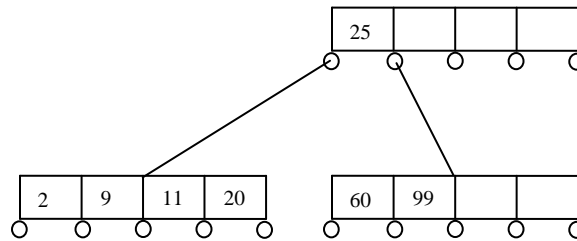
Suponer que ahora se desea dar de baja el nodo 23. En este caso el lugar que deja el 23 es ocupado por el elemento que está más a la izquierda del hijo derecho como se muestra.



Suponer que ahora se desea dar de baja el elemento 18. Al momento de darlo de baja se tiene la siguiente información.

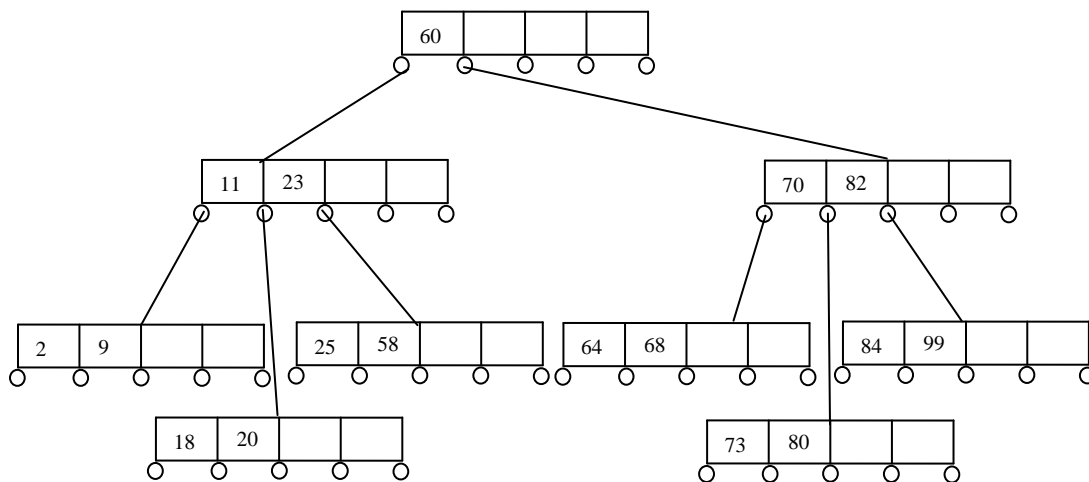


Pero no es válido que un nodo tenga menos de (entero  $(m/2)$ ) lugares ocupados. Lo primero que se tiene que hacer es ver si el hermano izquierdo o el derecho tiene más del mínimo permitido. Como no es así; ya que el mínimo en un árbol de orden cuatro es (entero  $(4/2)=2$ ), se fusionan el hijo izquierdo el padre y el hijo derecho en un solo nodo como se muestra a continuación.

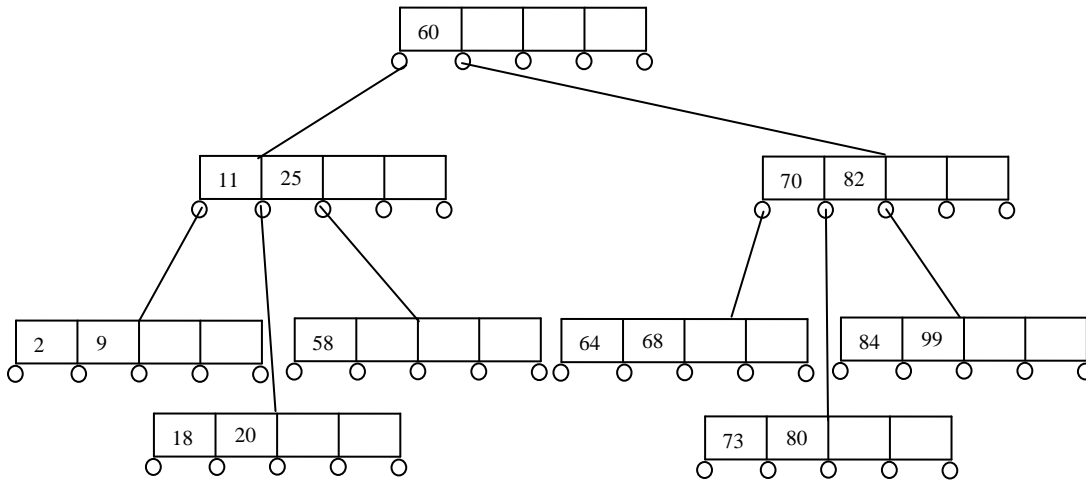


Ahora se debe comprobar que el nodo padre tenga el mínimo de datos exigidos, pero como se trata de la raíz en este caso se permite que tenga menos de (entero ( $m/2$ )).

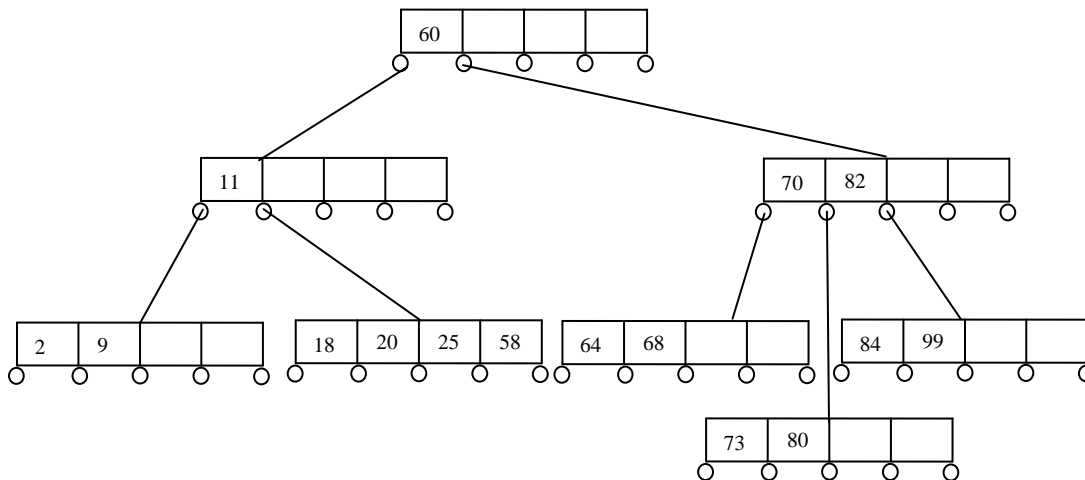
Considerar ahora; que se tiene el siguiente árbol B y se desea dar de baja el elemento 23.



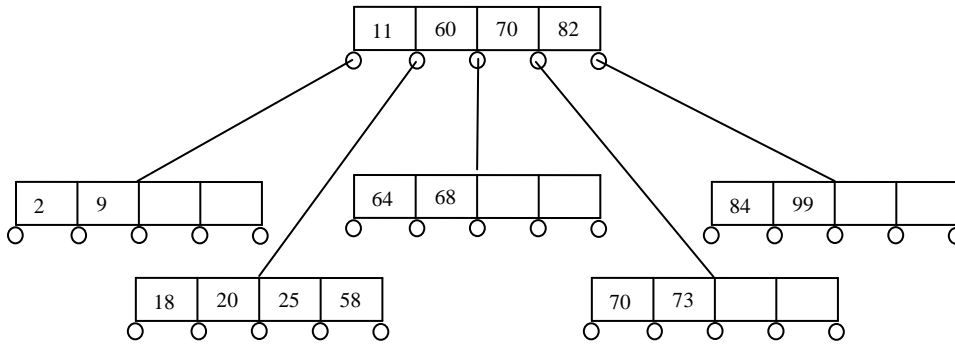
Lo primero que se hace es llevar a cabo un intercambio de lugares entre el 23 y el dato inmediatamente superior a 23 (en este caso el 25). Posteriormente borrar el 23 como se muestra en el siguiente árbol B.



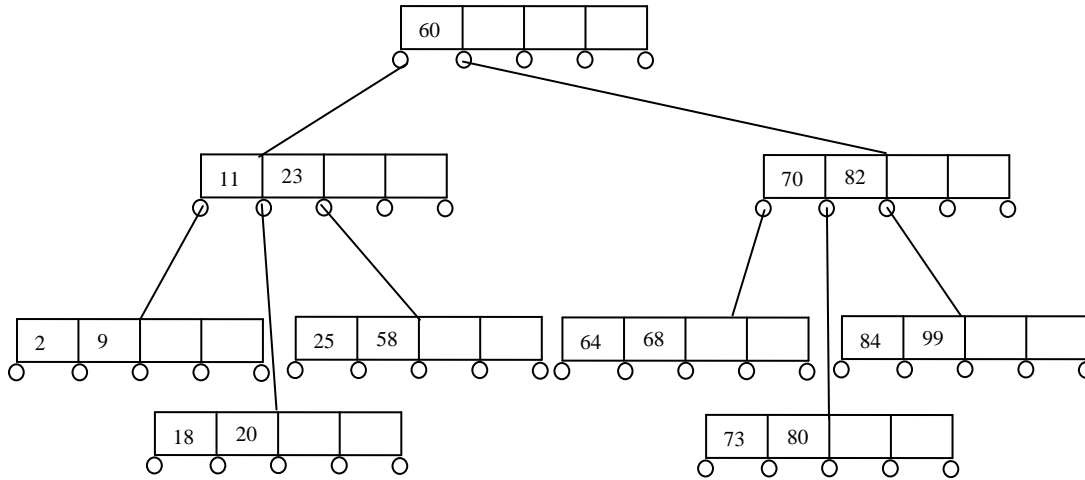
Cómo el nodo donde está el 58 no cumple con el mínimo requerido de datos se fusionan el hijo izquierdo (18, 20) con el nodo padre (25) y el hijo derecho (58), para obtener el siguiente árbol.



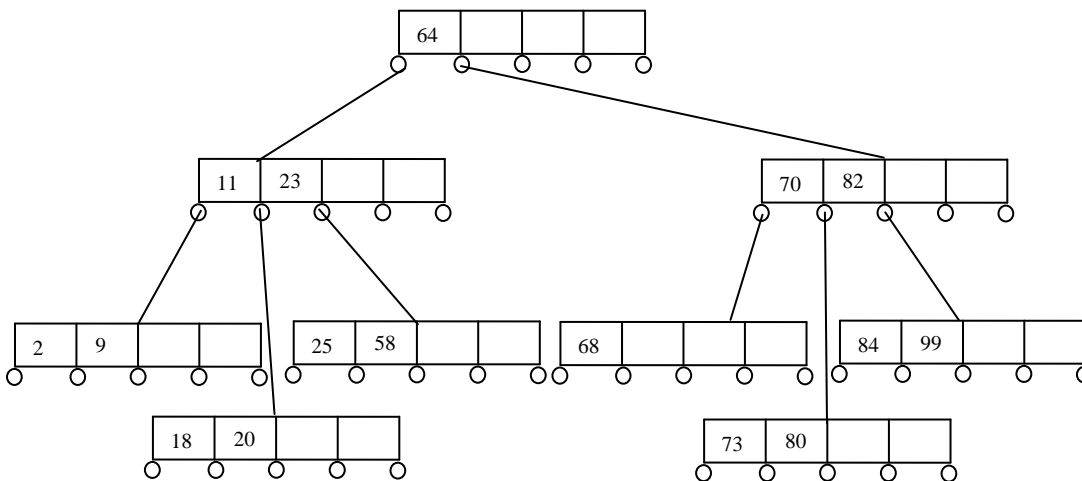
Pero nuevamente el nodo donde está el dato 11 no tiene los elementos mínimos requeridos, por lo tanto se debe fusionar el hijo izquierdo (11) el padre (60) y el hijo derecho (70, 82) para obtener el siguiente árbol.



Suponer ahora que se tiene el siguiente árbol B y que se desea dar de baja el elemento 60 que está en el nodo raíz.

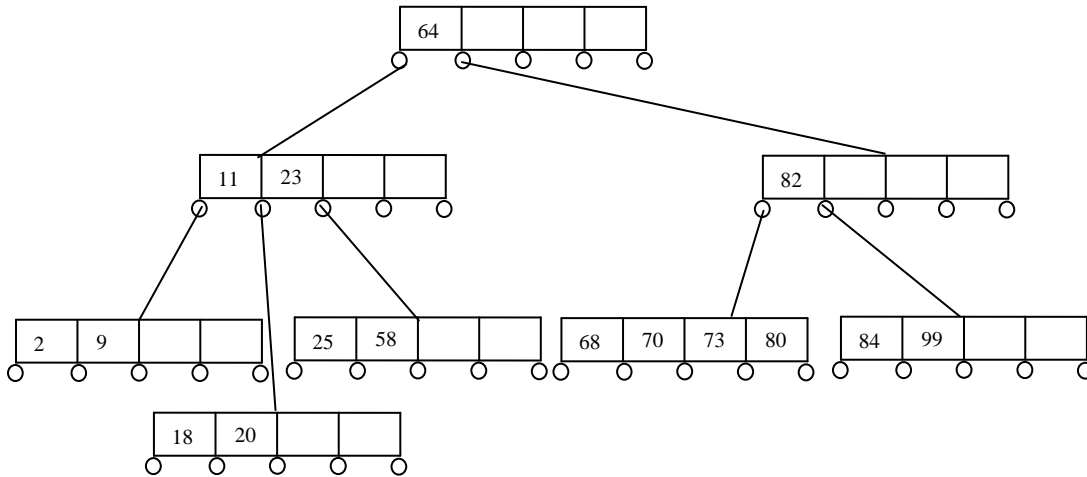


Lo primero es intercambiar de lugares el elemento 60 con el inmediato superior en este caso el 64 y eliminar el 60 para obtener el siguiente árbol.

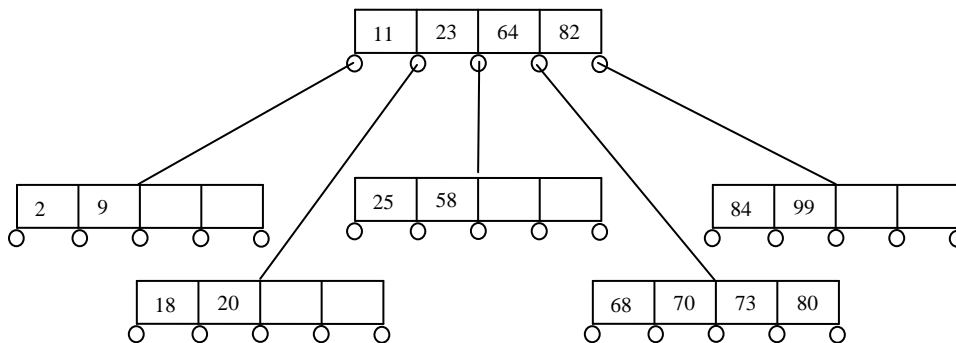




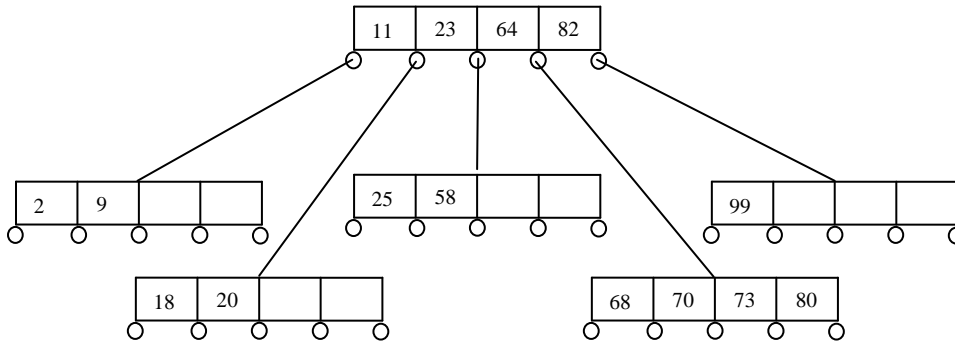
Pero el nodo donde está el 68 no tiene los elementos mínimos, por lo tanto se deben fusionar el hijo izquierdo (68) el padre (70) y el hijo derecho (73, 80) para obtener el siguiente árbol.



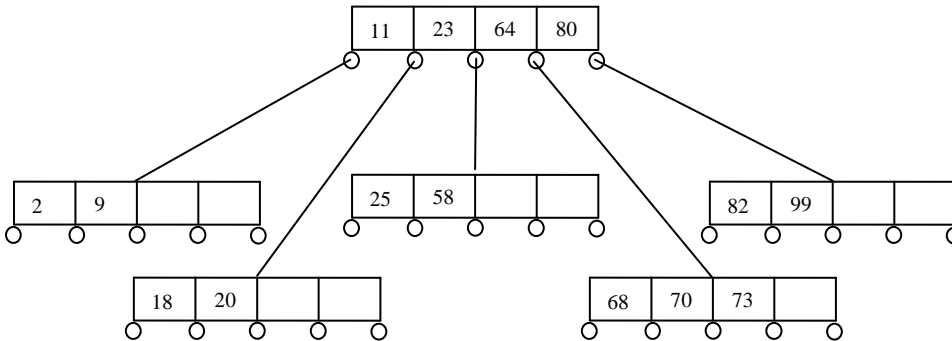
Nuevamente el nodo donde esta el elemento 82 no tiene los datos suficientes y se debe fusionar con el hijo izquierdo (11, 23) padre (64) e hijo derecho (82) resultando el árbol.



Finalmente suponer que en este árbol resultante se desea borrar el elemento 84, de tal forma que el árbol queda de la siguiente manera.



Pero el nodo en donde está 99 no tiene el mínimo número de elementos requeridos. Se fusionan el hijo izquierdo (68, 70, 73, 80) con el padre (82) y el hijo derecho (99) resultando: (68, 70, 73, 80, 82, 99). En general se toma como padre al intermedio del conjunto, pero como en este caso no es único, ya que son intermedios 73 y 80, por lo general se toma el mayor de ellos (80), resultando el siguiente árbol B.



Por todo lo anteriormente visto en este tema de árboles, se puede decir que dependiendo del fin perseguido, el volumen y características de la información, así deberá ser el árbol que se debe usar para obtener los mejores resultados.

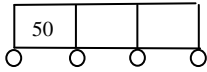
Ejemplo:

Sea el siguiente conjunto de datos: 50, 30, 65, 42, 18, 14, 70, 79, 69, 22, 16.

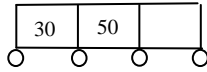
- Crear un árbol B de orden 3 (mostrar la secuencia).
- Dar de baja los elementos 42, 70 y 50 (Mostrar secuencia en cada caso).
- En este árbol cuál es el número:
  - Máximo de datos por nodo.
  - Mínimo de datos por nodo.
  - Máximo de ramas por nodo.
  - Mínimo de ramas no nulas por nodo.

Respuestas:

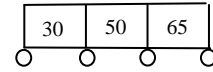
a) Creación del Árbol B de orden 3.



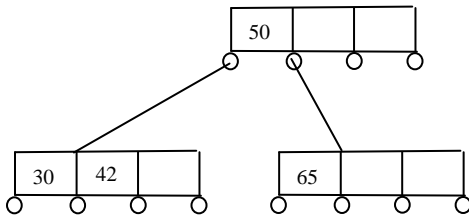
Alta 30



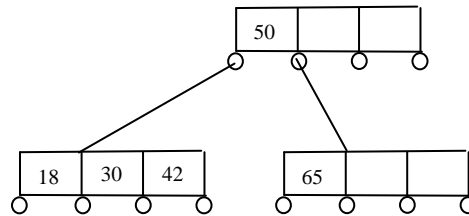
Alta 50



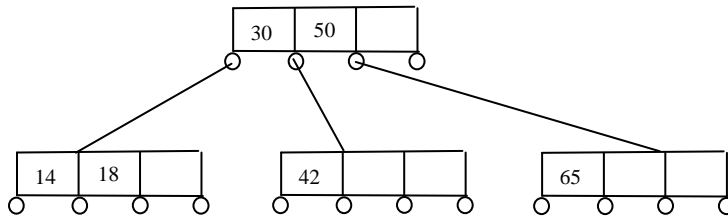
Alta 65



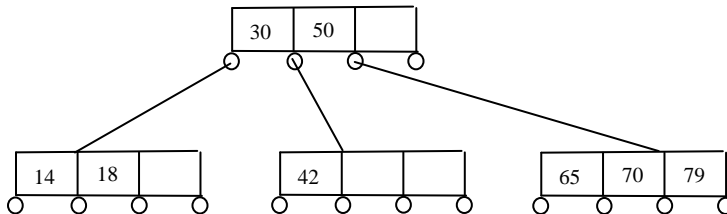
Después de dar de alta a 42 y tomar el mayor de los datos intermedios (50) como padre.



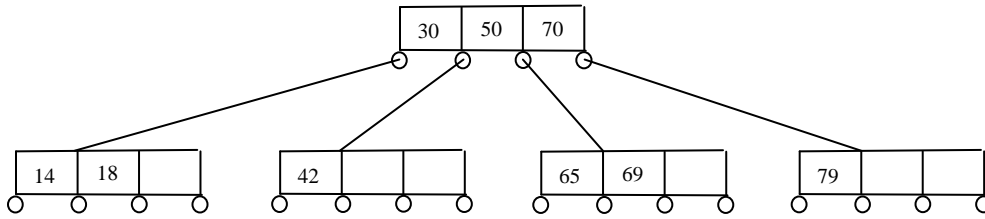
Alta 18.



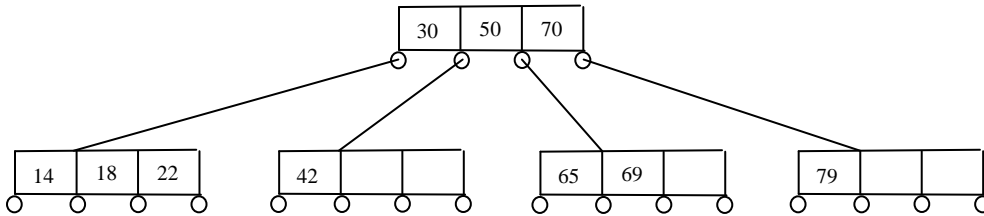
Después de dar de alta al 14 y tomar el mayor de los datos intermedios (30) como padre.



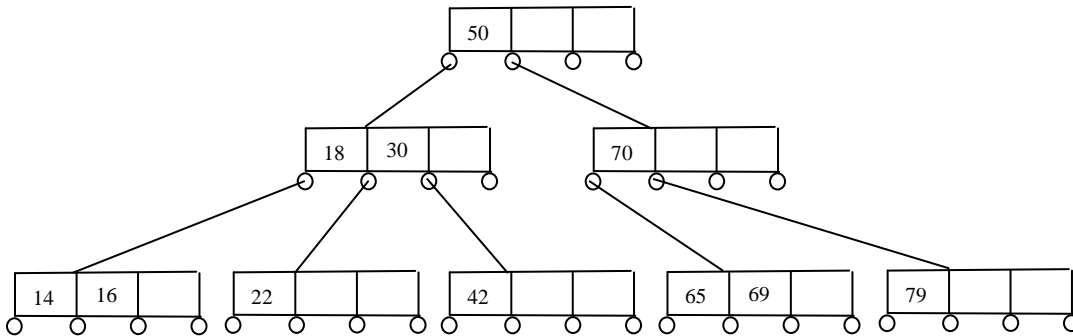
Alta 70 y 79



Después de dar de alta el 69 y tomar el mayor de los datos intermedios (70) como padre.

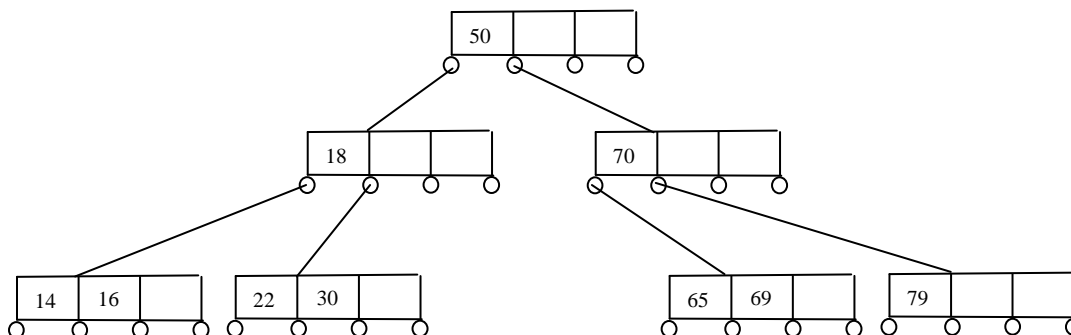


Alta 22

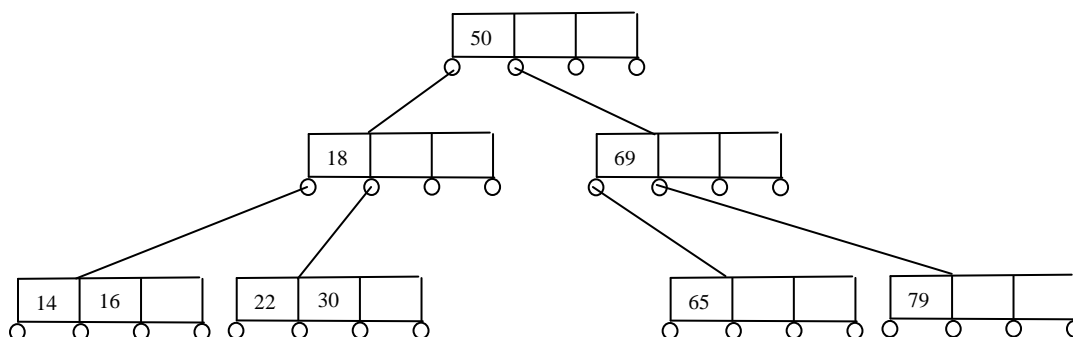


Después de dar de alta el 16 y tomar el mayor de los datos intermedios (18) como padre, y posteriormente tomar el mayor de los intermedios (50) como padre.

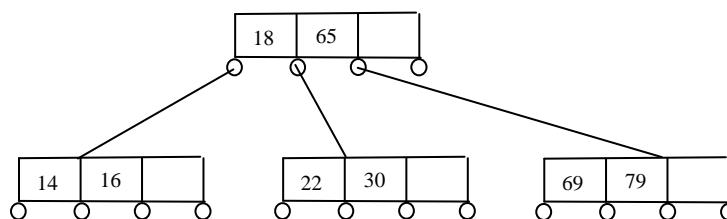
b) Bajas de los elementos 42, 70 y 50.



Después de dar de baja el 42 .



Después de dar de baja el 70 .



Después de dar de baja el 50

- c) En este árbol el número:
- Máximo de datos por nodo es el orden  $m$  (en este caso  $m=3$ ).
  - Máximo de ramas ( $n= m+1$ ) por nodo (en este caso  $n = m+1 = 3+1$ )
  - Mínimo de datos por nodo (entero ( $m/2$ )). En este caso  $\text{entero}(m/2) = \text{entero}(3/2) = 1$
  - Mínimo de ramas no nulas por nodo ( $\text{entero}(n/2+0.5)$ ). En este caso  $\text{entero}(n/2+0.5) = \text{entero}(4/2+0.5) = 2$ .

Ejercicios:

Crear un árbol B y llevar a cabo las bajas que se indican con la información que se proporciona en cada uno de los incisos. Si hay elementos repetidos, considerarlos como menores en la segunda aparición.

- a) Orden 3. Conjunto de elementos iniciales: 13, 20, 15, -8, 0, 40, 69, 12, 58, 23, 21, 33. Elementos a dar de baja: 20, 23 y 69.
- b) Orden 4. Conjunto de elementos iniciales: 90, 12, -7, 24, 0, 72, 15, 68, -3, 14, 53, 28, -9, 3, 61, 52, 27. Elementos a dar de baja: 14, 15 y -3.
- c) Orden 5. Conjunto de elementos iniciales: 2, 6, 10, 4, 8, 1, 7, 67, 14, 81, 33, 21, 58, 43, 68, 12, 17, 45, 39, 21, 19, -1, 70, 40, 30, 23. Elementos a dar de baja: 40, 43, 45, 39, 33, y 10.
- d) Orden: 6. Conjunto de elementos iniciales: 90, 10, 80, 20, 70, 30, 60, 40, 50, 12, 45, 35, 75, 55, 95, 15, 25, 1, 2, 3, 4, 5, 6, 7, 92, 11, 19, 28, 48. Elementos a dar de baja: 80, 90, 70, 92, 60 y 95.

Mostrar el procedimiento en cada uno de los incisos y también para cada inciso encontrar el número:

- Máximo de datos por nodo.
- Mínimo de datos por nodo.
- Máximo de ramas por nodo.
- Mínimo de ramas no nulas por nodo.