

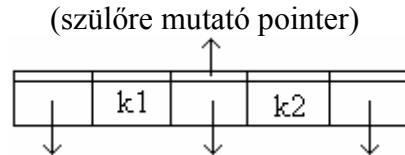
17. A 2-3 fák és B-fák

2-3 fá

Fontos jelentősége, hogy belőlük fejlődtek ki a B-fák.

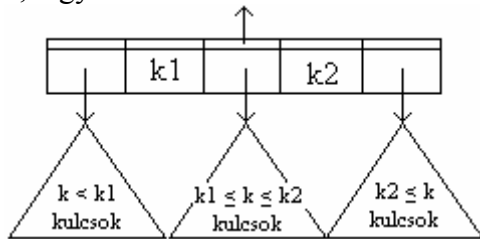
Def.: Minden belső csúcsnak 2 vagy 3 gyermeke van. A levelek egy szinten helyezkednek el. Az adatrekordok/kulcsok csak a levelekben vannak tárolva. A belső pontokban a keresést informáló kulcsok és mutatók vannak. A levelekben balról jobbra nőnek a kulcsok.

A belső pontokban fizikailag ilyen szerkezet van:



Logikailag a belső csúcs kétféle:

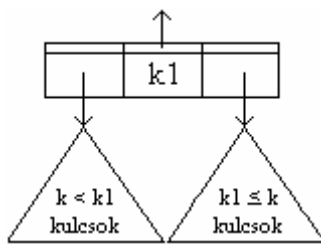
a, 3 gyerek



[+] k1 a legkisebb kulcs k2 a legkisebb kulcs

Ez az előírás nem mindenütt szerepel, nálunk igen!

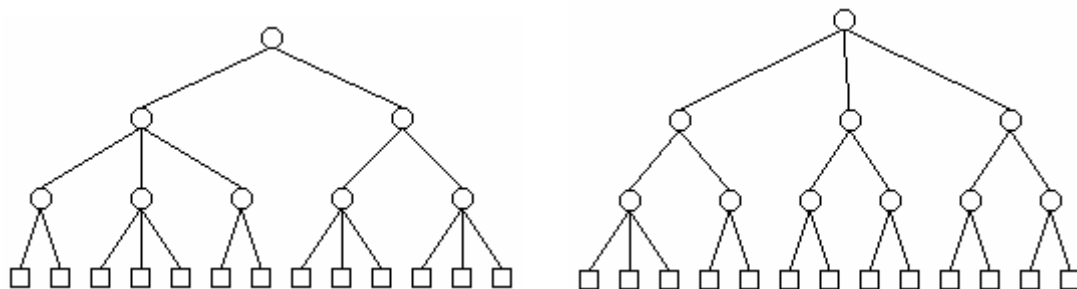
b, 2 gyerek



[+] k1 a legkisebb kulcs

A két típus közötti váltást csak logikailag kezeljük. („csere”)

Pl.: 2 különböző szerkezetű 2-3 fa, n = 13



Megjegyzés: $n = 0$ $t = \text{NIL}$ vagy üres gyökér
 $n = 1$ kivételesen a gyökérnek egy gyereke van.

Összefüggés n és h között:

$$2^h \leq n \leq 3^h$$

$$h \leq \log_2 n$$

Még 2-es elágazási tényező esetén is korlátos a fa magassága.

Műveletek:

Keresés

Összehasonlítások száma $1, 2, \dots, h-1$ magasságban: $1-2$

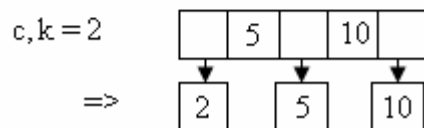
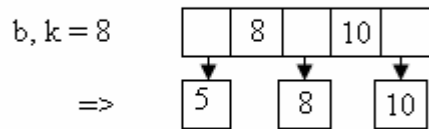
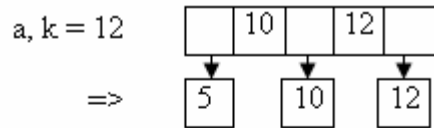
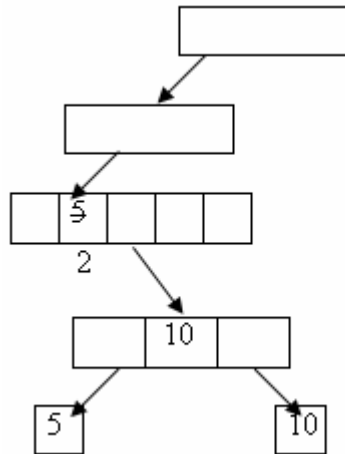
h magasságban: 1

$$T(n) \leq 2h+1 \leq 2\log_2 n+1 = O(\log n)$$

Beszűrés

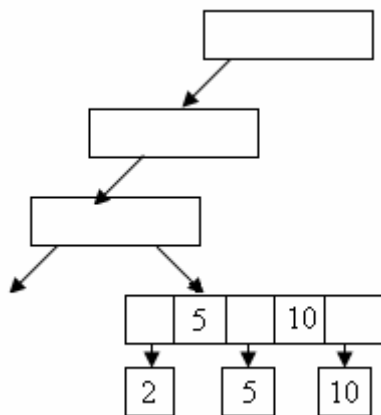
Kereséssel meghatározzuk a helyét.

(I) A legelső belső pontnak 2 gyereke van (elfér még egy harmadik is)

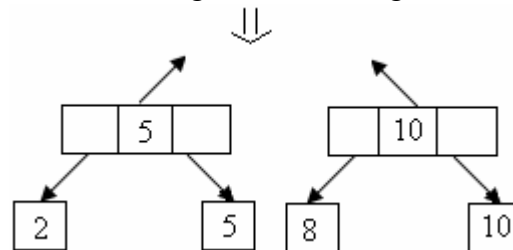


Továbbá felfelé haladva korrigálni kell a megfelelő kulcsértéket 2-re.

(II) A legelső belső pontnak 3 gyereke van



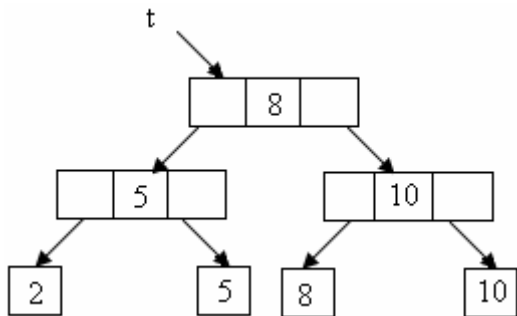
$k = 8$ A megoldás a csúcsvágás.



A szülőben is módosítani kell.

Ha a szülőnek eleve 3 gyereke volt, akkor itt is csúcsvágásra van szükség, és így tovább felelé. Ha valahol ezen az úton van egy 2 gyermekes belső csúcs, akkor ott megáll a beszűrés, mert annak lehet 3 gyereke.

Ha ezen az úton minden belső pontnak 3 gyereke volt, akkor a csúcsvágás felgyűrűzik a gyökérig. Ekkor a kétfelé vágott gyökér fölé egy új gyökeret kell tenni. Az előző ábrát véve alapul:

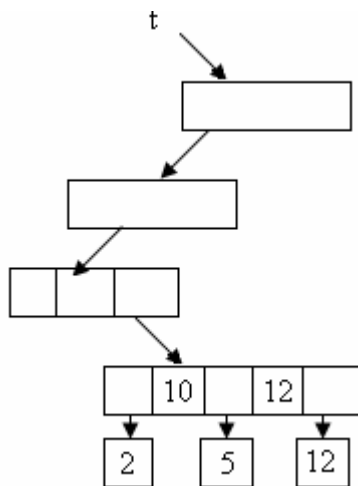


Nőtt a h!

Törlés

Megkeressük a törlendő kulcsot.

(I) A kulcs szülőjének 3 gyereke van (tehát neki 2 testvére van).



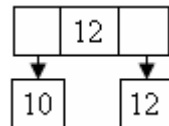
a, $k = 5$

b, $k = 10$

c, $k = 12$

a,

=>

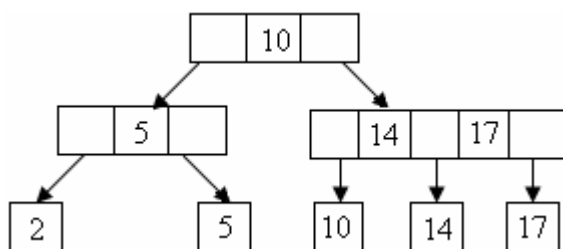


+ korrekció a szülőben

$5 \Rightarrow 10$ -esre

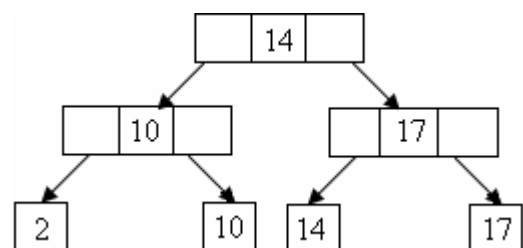
(II) A törlendő elemnek csak 1 testvére van, azaz a szülőnek 2 gyereke van.

II / 1 Ha a szülőnek van 3 gyerekes testvére, akkor az 1 gyereket átad

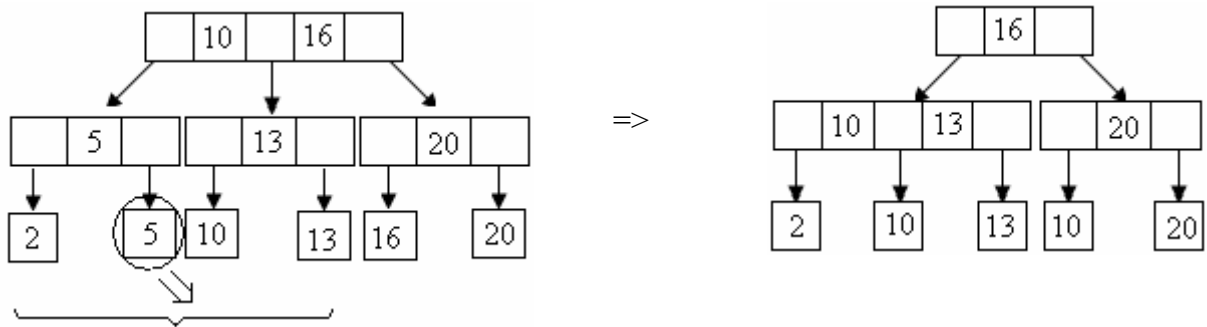


$k = 5$

=>



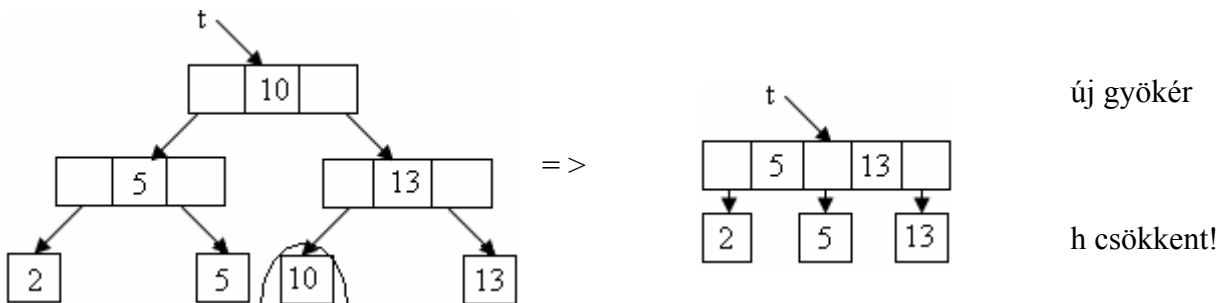
II / 2 Ha a szülőnek nincs 3 gyerekes testvére



csúcsösszevonás

Szükség esetén a csúcsösszevonásokat folytatjuk felfelé. Ez felgyűrűzhet a gyökérig.

k = 10



új gyökér

h csökkent!

Műveletek költsége:

$$T(n) = O(h) = O(\log n)$$

Megjegyzés: Invariáns (+) feltétel akkor sérül, amikor a minimális elemet töröljük, vagy a minimálissal kisebb elemet szúrunk be.

B-fák

A 2-3 fák természetes általánosításai: $3 \Rightarrow r$, $2 \Rightarrow \left\lceil \frac{r}{2} \right\rceil$; gyakorlatban kb. $50 \leq r \leq 1000$ (egy

fára r rögzített!)

A háttértárolókon való nagy mennyiségű adat tárolásának legjobb módszere.

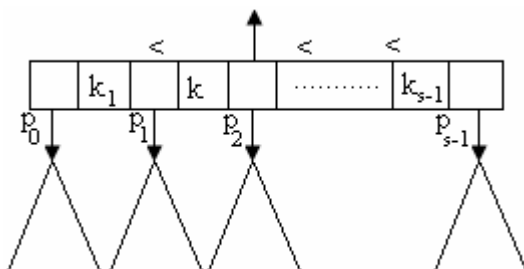
Elméleti alapok:

Definíció: B-fa (r-áris / r-edrendű kereső-/rendezőfa):

A gyökér s leágazásszáma $2 \leq s \leq r$ (kivéve az $n=0,1$ eseteket)

A (gyökér alatti) többi belső pont száma: $\left\lceil \frac{r}{2} \right\rceil \leq s \leq r$

Egy belső pont logikai tartalma:



Az itteni k kulcsokra: $k_1 \leq k < k_2$ és k_1 minimális kulcs.

A levelek mind egy szinten, a gyökértől egyenlő távolságra helyezkednek el.

A tulajdonképpeni adatok (rekordok) itt vannak.

A teljes fa esetleg néhány felső szint kivételével háttértárolón van. A levelek ilyenkor általában több rekordot tartalmaznak.

Műveletek:

Keresés, beszúrás és törlés: a keresőfák ill. 2-3 fák alapján könnyen elképzelhető.

Meggondoljuk:

Keresés:

A belső pontokban lineáris vagy logaritmikus keresés; lineáris keresés esetén $M\ddot{O} = r-1$

Még a levélben is keresünk több rekord esetén.

Beszúrás:

Itt is csúcsvágással (ha elértük r-et)

Pl.: $r = 11$, $\left\lceil \frac{r}{2} \right\rceil = 6$

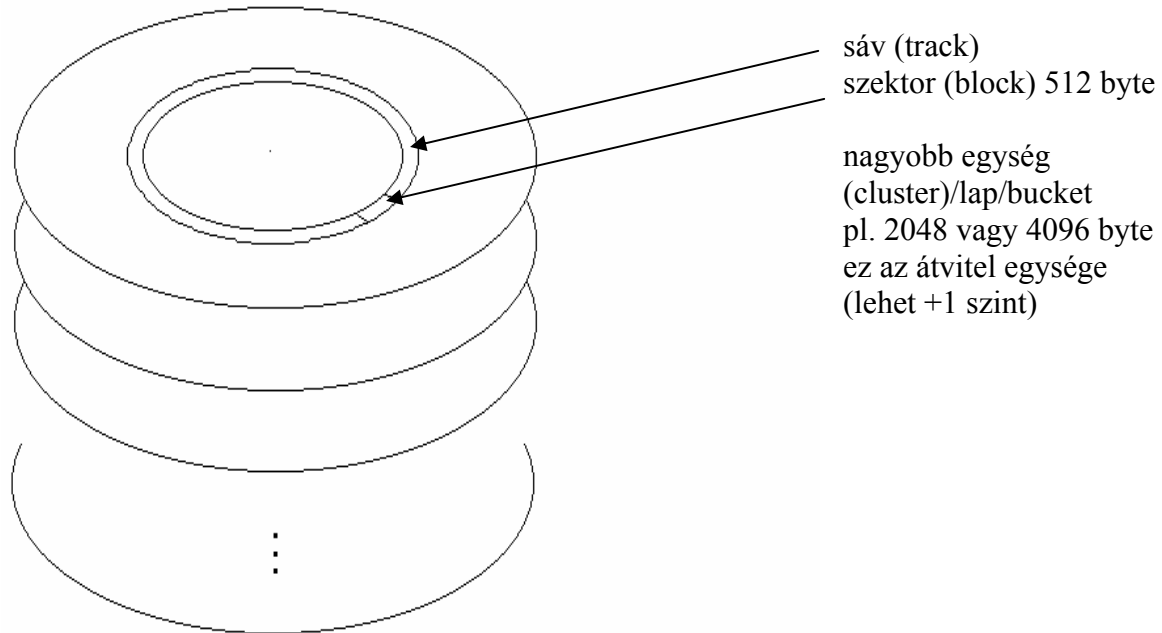
Ha egy betelt csúchhoz adunk +1 gyereket, akkor 12 lenne, amiből előállítunk 6+6-ot.

Törlés:

Itt is csúcsösszevonással (ha elértük az $\left\lceil \frac{r}{2} \right\rceil$ -t lefelé)

Ha egy pont alatt 6 gyerek van (előző pont: Beszúrás –beli példánknál) és ennek minden testvére alatt is, akkor a törlés után 5 lenne $\Rightarrow 5+6=11$

Mágneslemezes háttértárolón való elhelyezés



„lap” egy olyan logikai egység, amelyet a rendszer együtt kezel (átvitel)

Legyen pl. 1000000 rekord

1 lap = 2048 byte

1 rek = csaknem 500 byte (480)



1 lapon 4 rekord fér el.

kulcs 16 byte

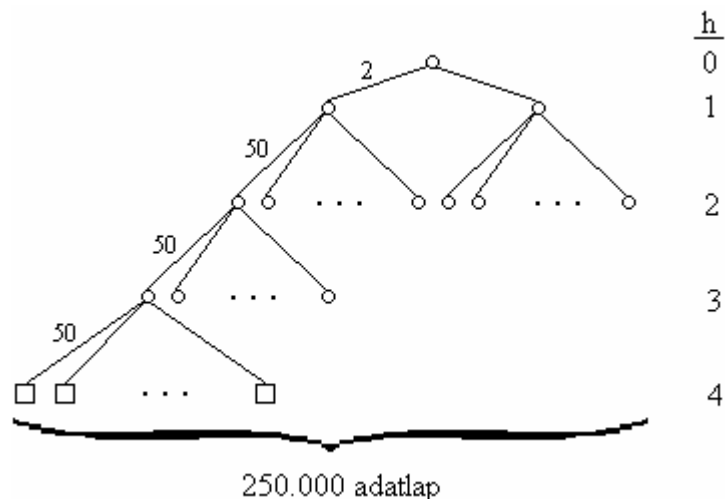
címek 4 byte

Ekkor egy belső pont (= index lap) esetén $r = 100$

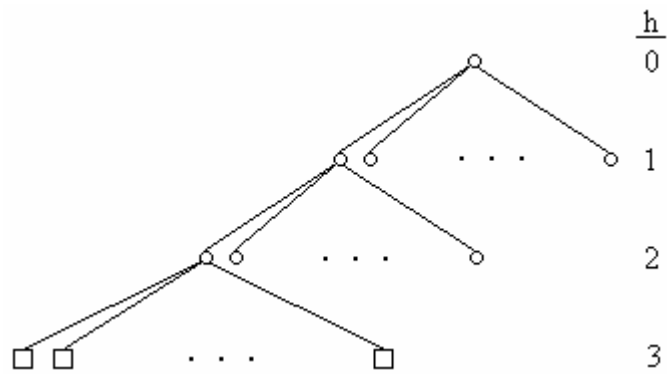
A levelek száma = 250000

A fa mérete:

max h



min h



Az ábrázolás paramétereinek a kritériuma:

Az átlagos elérése idő ~ lemezhez fordulások száma.

Itt: 3-4

Fontos javítás: gyökér mindig a memóriában van, esetleg még az alatta lévő szint is; kiegészítő memóriaterületen [cash] bizonyos elérési útvonalak tárolása.