Algoritmusok és adatszerkezetek II.

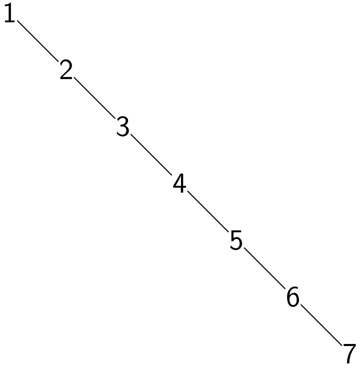
Piros-fekete fák

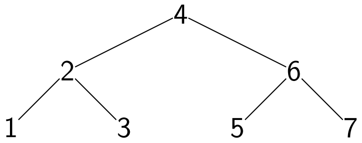
Szegedi Tudományegyetem

Mit értünk kiegyensúlyozott keresőfa alatt?

Emlékeztető

Az eddig tárgyalt műveletek h magas fákra $O(h)$ idejűek voltak
A továbbiakban szeretnénk, ha $h = \Theta(\log n)$ is teljesülne





Piros-fekete fák tulajdonságai

1. Minden csúcs színe piros vagy fekete
2. A gyökér színe fekete
3. Minden levele[1] fekete
4. A piros csúcsoknak **kizárólag** fekete színű gyerekeik vannak
5. Bármely csúcsból azonos számú fekete csúcs érintésével jutunk el bármelyik levélbe

[1] levelek alatt itt most az "őrszemeket" értjük

Piros-fekete fák tulajdonságai

1. Minden csúcs színe piros vagy fekete
2. A gyökér színe fekete
3. Minden levele[1] fekete
4. A piros csúcsoknak **kizárólag** fekete színű gyerekeik vannak
5. Bármely csúcsból azonos számú fekete csúcs érintésével jutunk el bármelyik levélbe

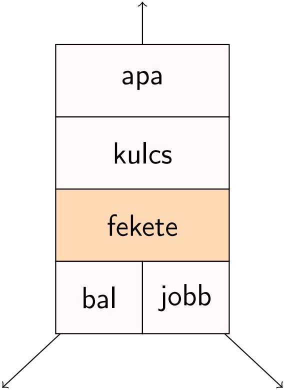
Tétel

Bármely n kulcsú piros-fekete fa magassága legfeljebb $2\log(n+1)$.

[1] levelek alatt itt most az "őrszemeket" értjük

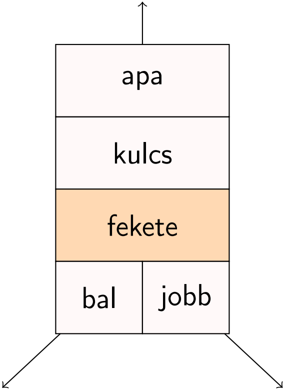
Piros-fekete fa implementációja

```
class Node {
    Object kulcs;
    boolean fekete;
    Node *apa;
    Node *bal;
    Node *jobb;
}
```



Piros-fekete fa implementációja

```
class Node {
    Object kulcs;
    boolean fekete;
    Node *apa;
    Node *bal;
    Node *jobb;
}
```



Megjegyzés

Az eddigi kiegészítő információk közül a legolcsóbb (csupán 1 bit)

Legalább hány kulcsból áll egy h magas AVL fa?

magasság	m
1	1
2	2
3	1+2+1=4
4	2+4+1=7
5	4+7+1=12
6	7+12+1=20
7	12+20+1=33
8	20+33+1=54
$\vdots$	$\vdots$

Legalább hány kulcsból áll egy h magas AVL fa?

magasság	m
1	1
2	2
3	1+2+1=4
4	2+4+1=7
5	4+7+1=12
6	7+12+1=20
7	12+20+1=33
8	20+33+1=54
$\vdots$	$\vdots$
h	$m_{h-2} + m_{h-1} + 1$

Miért igaz, hogy az AVL fák $O(\log n)$ magasak?

- Jelölje m_h a h magas AVL fában lévő minimálisan található kulcsok számát ($m_1 = 1, m_2 = 2$)
- Általánosságban ($h > 2$ esetén): $m_h = m_{h-2} + m_{h-1} + 1$
- $m_{h-2} < m_{h-1}$ természetesen teljesül, ahonnan

$$m_h > 2m_{h-2}$$

Miért igaz, hogy az AVL fák $O(\log n)$ magasak?

- Jelölje m_h a h magas AVL fában lévő minimálisan található kulcsok számát ($m_1 = 1, m_2 = 2$)
- Általánosságban ($h > 2$ esetén): $m_h = m_{h-2} + m_{h-1} + 1$
- $m_{h-2} < m_{h-1}$ természetesen teljesül, ahonnan

$$m_h > 2m_{h-2} > 2 * 2m_{h-4} > \ldots > 2^i m_{h-2i}$$

Miért igaz, hogy az AVL fák $O(\log n)$ magasak?

- Jelölje m_h a h magas AVL fában lévő minimálisan található kulcsok számát ($m_1 = 1, m_2 = 2$)
- Általánosságban ($h > 2$ esetén): $m_h = m_{h-2} + m_{h-1} + 1$
- $m_{h-2} < m_{h-1}$ természetesen teljesül, ahonnan

$$m_h > 2m_{h-2} > 2 * 2m_{h-4} > \ldots > 2^i m_{h-2i}$$

- $m_h > 2^i m_{h-2i}$ összefüggést m_1-ig kijátszva $m_h > 2^{h/2}$

Miért igaz, hogy az AVL fák $O(\log n)$ magasak?

- Jelölje m_h a h magas AVL fában lévő minimálisan található kulcsok számát ($m_1 = 1, m_2 = 2$)
- Általánosságban ($h > 2$ esetén): $m_h = m_{h-2} + m_{h-1} + 1$
- $m_{h-2} < m_{h-1}$ természetesen teljesül, ahonnan

$$m_h > 2m_{h-2} > 2 * 2m_{h-4} > \ldots > 2^i m_{h-2i}$$

- $m_h > 2^i m_{h-2i}$ összefüggést m_1-ig kijátszva $m_h > 2^{h/2}$
- Tegyük fel, hogy egy h magas AVL fa $n \geq m_h$ csúcsból áll, azaz $n > 2^{h/2}$, vagyis $h < 2\log_2(n)$

Miért igaz, hogy az AVL fák $O(\log n)$ magasak?

- Jelölje m_h a h magas AVL fában lévő minimálisan található kulcsok számát ($m_1 = 1, m_2 = 2$)
- Általánosságban ($h > 2$ esetén): $m_h = m_{h-2} + m_{h-1} + 1$
- $m_{h-2} < m_{h-1}$ természetesen teljesül, ahonnan

$$m_h > 2m_{h-2} > 2 * 2m_{h-4} > \ldots > 2^i m_{h-2i}$$

- $m_h > 2^i m_{h-2i}$ összefüggést m_1-ig kijátszva $m_h > 2^{h/2}$
- Tegyük fel, hogy egy h magas AVL fa $n \geq m_h$ csúcsból áll, azaz $n > 2^{h/2}$, vagyis $h < 2\log_2(n)$

Megjegyzés

Az élesebb $h < 1.44\log_2(n)$ korlát is bizonyítható.

Összegzés

- A bináris keresőfák műveletei $O(h)$ idejűek
- Legrosszabb esetben azonban n is lehet a fák magassága ($\Theta(\log n)$ helyett)
- Kiegyensúlyozott keresőfák használatával garantálható, hogy a keresőfa kiegyensúlyozottsága sose romoljon el "túlságosan"