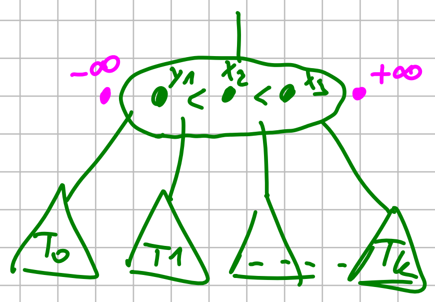
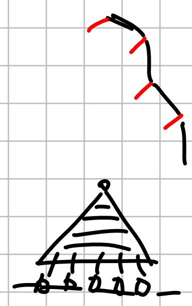
# Vícecestný vyhledávací strom




- zakořeněný strom
- má int. + ext. vrcholy
- synové vrcholu mají pořadí
- v int. vrcholech jsou klíče (aspoň 1)

$-\infty = x_0 < x_1 < \dots < x_k < x_{k+1} = +\infty$

- # synů = # klíčů + 1
  ↳ podstromy $T_0, \dots, T_k$
- uspořádání: $\forall i\ \forall y$ klíč v $T_i$:
  $x_i < y < x_{i+1}$

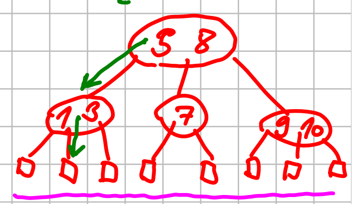


parametry: $a \geq 2$, $b \geq 2a-1$

**(a,b)-strom** je vícecestný VS t.ž.:

1. všechny ext. vrcholy jsou stejně hluboko
2. int. vrcholy mají $a$ až $b$ synů ($a-1$ až $b-1$ klíčů)
   výjimka: kořen má 2 až $b$ synů (1 až $b-1$ klíčů)

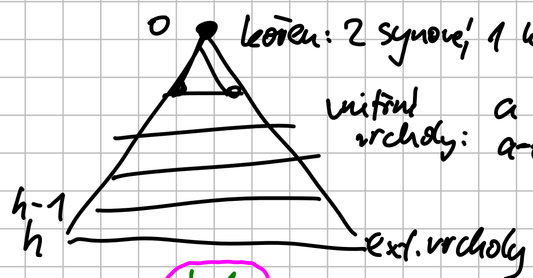
příklad: (2,3)-strom
2?



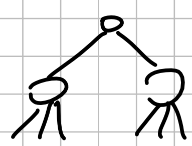

---

**Lemma:** (a,b)-strom s $n$ klíči má hloubku $\Theta(\log n)$.
← závisí na $a, b$

**Dk:** $m_h$ := min. # klíčů ve stromu hloubky $h$
⇔ ∀ vrchol má min. možný # synů, tedy i klíčů




# vrcholů na hladině:
$1$
$2a^0$
$2a$
$2a^2$
$\vdots$
$2a^{i-1}$



Součet geom. řady:
$$\sum_{j=0}^{k} a^j = \frac{a^{k+1}-1}{a-1}$$

$$m_h = 1 + \sum_{i=1}^{h-1} 2a^{i-1}(a-1) = 1 + 2(\cancel{a-1}) \sum_{j=0}^{h-2} a^j \qquad \frac{a^{h-1}-1}{\cancel{a-1}}$$

$$= 1 + 2\left(a^{h-1} - 1\right)$$
$$= 2a^{h-1} - 1$$
← roste exponenciálně
⇓
max. hloubka pro $n$ klíčů roste logaritmicky ($\sim \log_a n$)

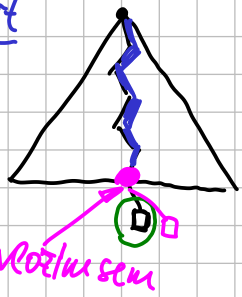
**Dolní mez:** $M_h$ := max. # klíčů ve stromě hloubky $h$

$M_h \sim b^h$ ⇒ min. hloubka roste též log. ($\sim \log_b n$)

---

**Find:** $O(1)$ na hladinu ⇒ $\Theta(\log n)$ celkem.

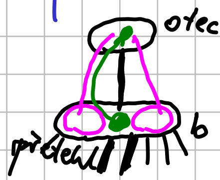
## Insert




Vložíme na nejnižší vnitřní hladinu.

Ošetříme přetečení (pokud nastalo):

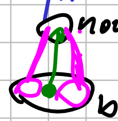
- předtím max. b-1 klíčů → při přetečení právě b



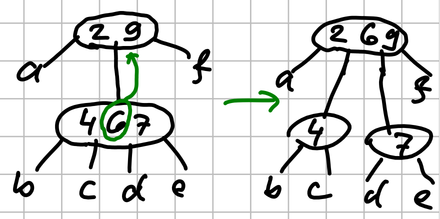

štěpení vrcholu

mohl přetéci otec

opakujeme .... nakonec můžeme štěpit kořen




(2,3)-strom 1 až 2 klíče/vrchol




Potenciální problém: poloviny příliš malé

pův. b klíčů, 1 jde do otce

poloviny mají $\lfloor (b-1)/2 \rfloor$ a $\lceil (b-1)/2 \rceil$ klíčů

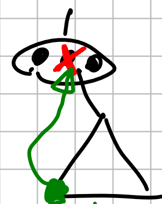
Problém jen pokud $\frac{b-1}{2} < a-1$

$b-1 < 2a-2$

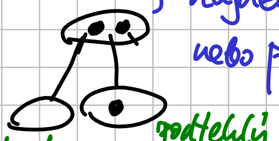
$b < 2a-1$ spor s definicí ($b \geq 2a-1$)

## Delete

Nejprve převedeme na Delete z nejnižší vnitřní hladiny

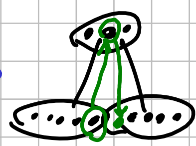



stačí umět řešit podtečení (vrchol s a-2 klíči)




najdeme levého (bližšího) nebo pravého bratra

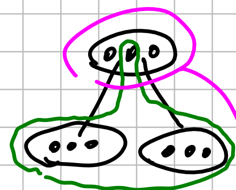
bratr má > a-1 klíčů



vyřešeno půjčením klíče

bratr má právě a-1 klíčů

slévame podteklý s bratrem



vyjde to:
$(a-2)+(a-1)+1 = 2a-2 \leq b-1$

otec mohl podtéci

opakujeme

→ při nejhorším podtekl kořen: je prázdný → smažeme ho

## Časová složitost:

$\Theta(1)$ času na hladinu, $\Theta(\log n)$ hladin

⇒ celkem $\Theta(\log n)$ na operaci.

## Volba a,b

- nechceme $b \gg 2a$ ... typicky buď $b=2a-1$ nebo $b=2a$
- nechceme velké a ... opt. jsou (2,3) nebo (2,4)

Na disku: 1 blok ~ 1 vrchol (a,b)-stromu

4 KB bloky, 4B klíče, 4B pointery

(256,512)-strom: 1 vrchol:

512 pointerů × 4B = 2KB

511 klíčů × 4B ≐ 2KB

4KB

strom o 3 vnitřních hladinách má aspoň $256^3 \doteq 16M$ klíčů

Paměť s keší ... bloky po 64B ← 1 vrchol má 1 blok

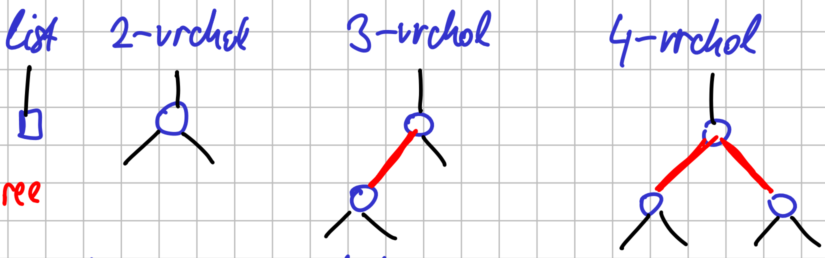
pro 4B klíče a 4B pointery : (4,8)-strom

**Na doma:**
- varianta: data jen v listech
- B-stromy ≈ (a,b)-stromy

---

**Překlad (2,4)-stromu na BVS**

Left-Leaning Red-Black Tree



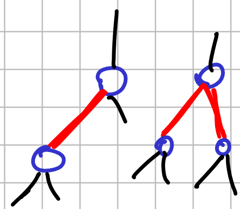

Ostrůvky

**Df:** LLRB strom je BVS s ext. vrcholy a hranami obarvenými červeně a černě t.ž.:
(pamatujeme si ve spodním vrcholu)

R axiomy:
1. nejsou 2R hrany těsně nad sebou
2. pokud z vrcholu dolů vede 1R hrana, pak doleva ] LL

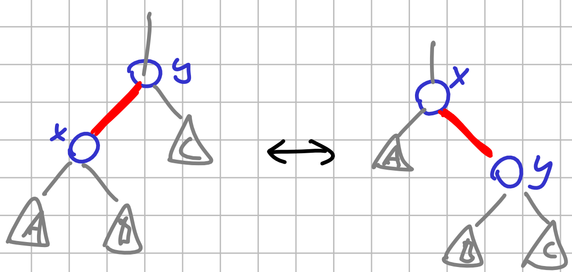
B axiomy:

3. hrany do listů jsou B
4. na ∀ cestě kořen-list je stejný #B hran



**Lemma:** ∃ bijekce mezi (2,4)-stromy a LLRB stromy.

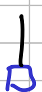
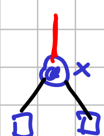
**Operace:** Rotace R hrany | Přebarvení 4-vrcholu




zachovává B axiomy, může rozbít R

zachovává B axiomy, může rozbít R

**Insert:** Směrem dolů přebarvujeme 4-vrcholy. → rozbíjí R

Nahradíme  za 


Směrem nahoru opravujeme R axiomy rotacemi.

→ trvá $\Theta(\log n)$