Datové struktury

rozhraní (statické / dynamické)

- fronta, zásobník, posloupnost
- prioritní fronta (halda)
- množina $X \subseteq \mathcal{U}$, $n = |X|$, $\mathcal{U}$ universum, X konečná množina
- slovník: částečné zobrazení $\{(k,v)\} \subseteq \mathcal{U}_k \times \mathcal{U}_v$ (k klíč unikátní, v hodnota)
- uspořádaná množina (slovník): $(\mathcal{U}, \leq)$ uspořádané universum

Find(x) = Member(x): zjistí, zda $x \in X$

Insert(x), Delete(x) … co když $x \in X$?

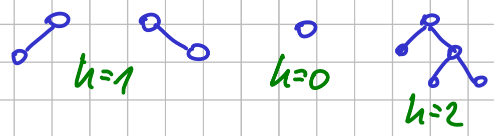
Min, Max, Pred(x) předchůdce, Succ(x) následník

Find	Insert	Delete	
$\Theta(n)$	$\Theta(1)/\Theta(n)$	$\Theta(n)$	pole
$\Theta(n)$	$\Theta(1)/\Theta(n)$	$\Theta(n)$	seznam
$\Theta(\log n)$	$\Theta(\log n)$	$\Theta(\log n)$	vyhl. strom
průměrně $\Theta(1)$	$\Theta(1)$	$\Theta(1)$	kešování

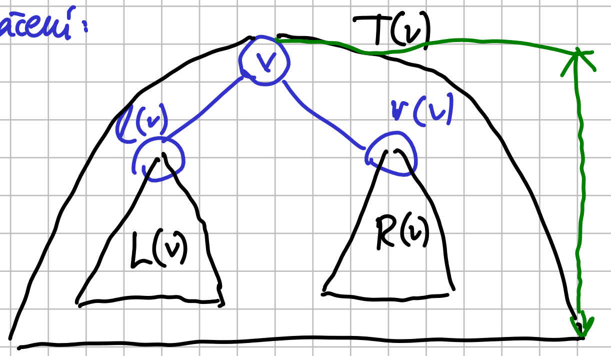
implementace: pole, spojový seznam, binární halda, binární vyhledávací strom

Binární strom

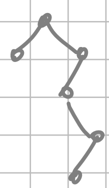
zakořeněný



Značení:



$h(v)$ hloubka podstromu $T(v)$ = max. # hran cesty mezi v a listem



n := celkový # vrcholů

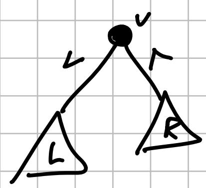
pokud chybí syn: $l(v)$ nebo $r(v) = \emptyset$

dodef. $T(\emptyset) = \emptyset$ $\quad h(\emptyset) = -1$

BVS (bin. vyhledávací strom) (a.k.a. BTS)

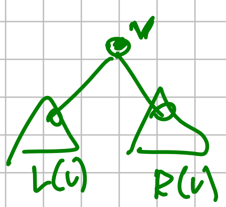
Vrcholy obsahují klíče $k(v) \in \mathcal{U}$

A navíc platí



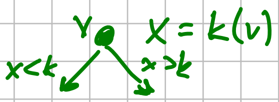
$\forall v$
$\forall \ell \in L(v)\ k(\ell) < k(v)$
$\forall r \in R(v)\ k(r) > k(v)$
$\Downarrow$
všechny klíče různé

Show (Enumerate)



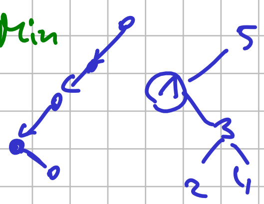
$\Theta(n)$

Find(x)

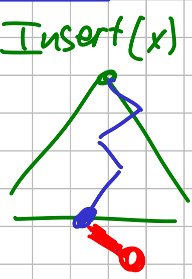


Θ(hloubka stromu)

Min

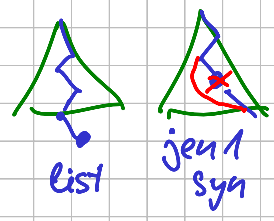


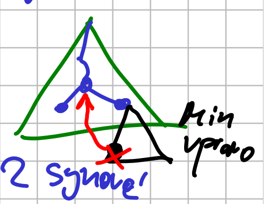
Insert(x)



Θ(hloubka)

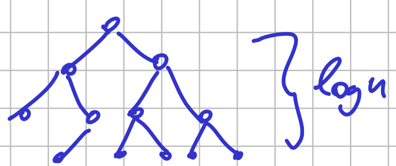
Delete(x)

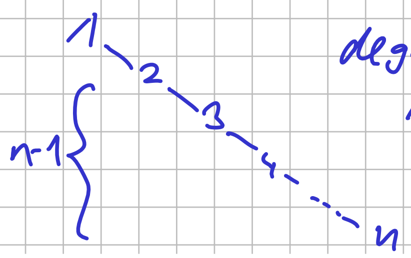




Θ(hloubka)

Chceme nízké stromy?





degenerovaný "liána"

Dokonale vyvážený BVS

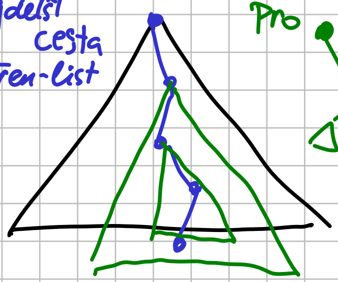
$n=2$ 

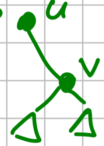
Df: Strom je dokonale vyvážený $\equiv$
$\forall v$ ~~$|L(v)| = |R(v)|$~~
$\big||L(v)| - |R(v)|\big| \le 1.$

Hloubka d.v. BVS $\le \log_2 n$

nejdelší cesta kořen-list

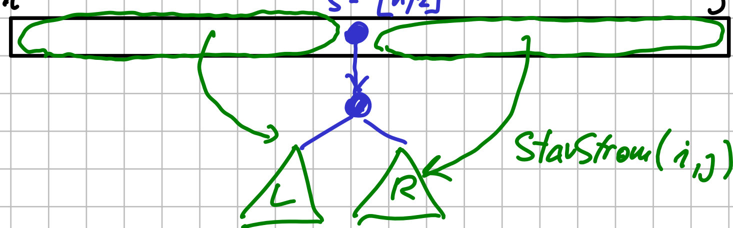
pro u, v platí $|T(v)| \le \frac{|T(u)|}{2}$





Posloupnost $x_1 < x_2 < \_\_\_ < x_n$
$\downarrow$
dokonale vyvážený strom

i $\quad s = \lfloor n/2 \rfloor \quad$ j



StavStrom(i,j)

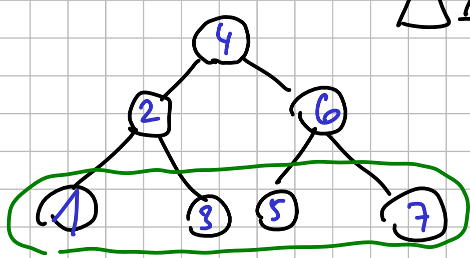
Rekurzivní alg. o složitosti $\Theta(n)$
$O(1)$ času na vrchol, n vrcholů

Věta: V každé implementaci operací Insert, Delete v d.v. BVS má aspoň 1 z operací složitost $\Omega(n)$ pro nekonečně mnoho hodnot n.

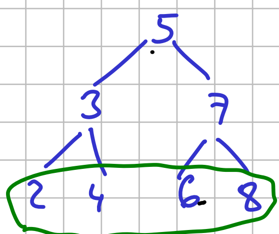
Dk: Zvolme $n = 2^k - 1$, klíče $1 \dots n$. $\longrightarrow$ jednoznačně určen strom

$n=7$

2^{k-1} $\quad \frac{2^k - 2}{2} = 2^{k-1} - 1$



lichá v listech



sudá v listech

provedeme Insert $(n+1)$, Delete (1) } potom $2 \dots n+1$ $\downarrow$ zase 1 značně

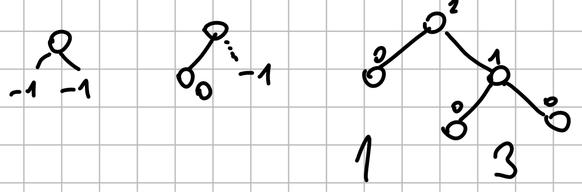
$\Omega(n)$ vrcholů změnilo, zda jsou listy
$\downarrow$
v každém změna min. 1 ukazatele
$\downarrow$
$\Omega(n)$ změn ukazatelů

pak ~~Insert $(n+2)$~~
Delete (2)
$\vdots$

za $\forall$ dvojici $\Omega(n)$ změn ukazatelů
$\hookrightarrow$ čas na dvojici je $\Omega(n)$
$\hookrightarrow$ Ins nebo Del trvá $\Omega(n)$

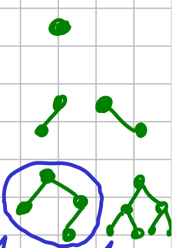
Df: Strom je hloubkově vyvážený $\equiv$
$\forall v \quad \big|h(l(v)) - h(r(v))\big| \le 1.$

AVL - stromy
1962



dokonale vyv. $\Downarrow$ $\not\Uparrow$ hloubkově vyv.

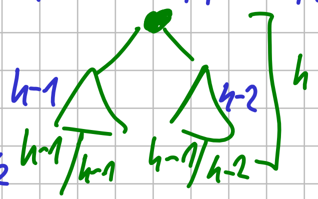
Věta: Hloubka AVL stromu s n vrcholy je $\Theta(\log n)$.



Dk: Počítáme $A_h :=$ min. # vrcholů AVL stromu hloubky h.

Dokážeme indukcí podle h, že $A_h \geq 2^{h/2}$.

$A_h = 1 + A_{h-1} + A_{h-2}$



(platí, že $A_h = F_{h+3} - 1$)

① pro $h=0$: $A_0 = 1 \geq 2^0 = 1$

pro $h=1$: $A_1 = 2 \geq 2^{\frac{1}{2}} = \sqrt{2} \doteq 1.414$

② pro $h \geq 2$: krok

$$A_h \geq A_{h-1} + A_{h-2} = 2^{h/2}\Big(\underbrace{\underbrace{\tfrac{\sqrt{2}}{2}}_{\doteq 0.707} + \underbrace{\tfrac{1}{2}}_{\doteq 0.5}}_{>1}\Big) \geq 2^{h/2}$$

podle IP:

$$A_{h-1} \geq 2^{\frac{h-1}{2}} = 2^{\frac{h}{2}} \cdot \underbrace{2^{-\frac{1}{2}}}_{\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}}$$

$$A_{h-2} \geq 2^{\frac{h-2}{2}} = 2^{\frac{h}{2}} \cdot \underbrace{2^{-1}}_{1/2}$$

$\Rightarrow \exists c > 1 \quad A_h \geq c^h$

$c = \sqrt{2} \doteq 1.414$

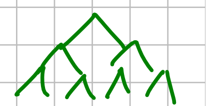
kdyby $h > \log_c n$

$A_h \geq c^{> \log_c n} > n$

$\Rightarrow$ strom na n vrcholech má hloubku $\leq \log_c n \in O(\log n)$.

[těsný odhad je $c = \frac{\sqrt{5}+1}{2} \doteq 1.618$]

Dolní odhad hloubky $\quad B_h :=$ max. # vrcholů stromu hloubky h



$B_h = 2^{h+1} - 1$ (indukcí)

$\leq 2^{h+1}$

$h \geq \log_2 n + 1 \quad \Rightarrow$ hloubka $\in \Omega(\log n)$.