5. cvičení

Datové struktury I, 1. 11. 2022 `https://iuuk.mff.cuni.cz/~chmel/2223/ds1/`

Úloha 1 (Hodnoty v listech vs vrcholech)
Na přednášce jste viděli variantu (a, b)-stromu s prvky v listech. Jak je potřeba strukturu a operace upravit, kdybychom měli prvky i ve vnitřních vrcholech?

Řešení
Stačí jít dle poznámek Martina Mareše – v principu jediný rozdíl je, že je potřeba poctivě migrovat hodnoty mezi vrcholy při štěpení/spojování vrcholů.

Úloha 2 ((a, b)-stromy na vlastní kůži)
Na obrázku (v zadání) máte $(2, 3)$-strom. Proveďte na něm následující operace (vždy jenom jednu, a poté začněte s novým stromem): Insert(7), Insert(48), Delete(44), Delete(40), Delete(32), Delete(30), Delete(16).

Řešení
Výsledky lze nasimulovat zde: `https://www.cs.usfca.edu/~galles/visualization/BTree.html`, původní strom vznikne přidáváním prvků v tomto pořadí: 28, 24, 26, 18, 0, 36, 16, 34, 22, 10, 2, 30, 4, 6, 20, 12, 32, 38, 14, 8, 40, 42, 44, 46

Úloha 3 (Správná volba parametrů)
Z přednášky víme, že libovolná posloupnost m insertů a deletů na $(a, 2a)$-strom celkem změní jenom $\mathcal{O}(m)$ vrcholů (když začínáme s prázdným stromem). Ukažte, že toto neplatí pro $(a, 2a-1)$-stromy, tedy pro libovolné m, n navrhněte posloupnost m operací na stromě s $\Theta(n)$ vrcholy, která celkem změní $\Omega(m \log n)$ vrcholů.
Můžete začít s (2,3)-stromy, a potom zobecnit pro libovolné a. Zároveň můžete začít s libovolným (validním) n-vrcholovým stromem, a až na konci ukázat, že jej opravdu vyrobit z prázdného stromu.

Řešení
Obecně: strom který má všechny vrcholy skoro prázdné (a synů), jen pravá cesta má všechny vrcholy plné ($2a - 1$ synů). Označíme-li si M jako 1+ největší číslo ve stromě, pak operace Insert(M),Delete(M) musí vždy všechno nejdříve rozštěpit, a poté spojit zpátky.
Po insertu totiž nejniží vrchol bude mít $2a$ synů, a tedy se bude muset rozštěpit na dva vrcholy s a syny - tím ale problém vybublá o úroveň výše, a opět musíme štěpit, až se takto dostaneme do kořene. Když pak provedeme delete, vrchol, který M obsahoval bude mít najednou jenom $a-1$ synů, a je „podměrečný“. Protože jeho jediný soused má také a synů, nemůžeme si žádného ukrást, a tedy musíme oba sloučit. Tím ovšem snížíme o 1 počet synů vrcholu o úroveň výše, a tohle nám opět probublá až do kořene.

Úloha 4 (Lepší zaplnění)
Zkuste upravit (a, b)-strom a jeho operace Insert, Delete tak, abychom mohli mít všechny vrcholy trochu plnější – konkrétně tak, abychom měli $\left(\frac{2}{3}b, b\right)$-strom.

Řešení
Trik je v tom vhodně spojovat (a rozdělovat) vrcholy - normálně spojujeme dva vrcholy do jednoho, když jsou oba malé.
Insert upravíme následovně: prvek vložíme jako obvykle, a pokud máme přeplněný vrchol, pokusíme se jeden z prvků přesunout do sourozence. Pokud je sourozenec také plný, tak potom tyto dva vrcholy, které mají dohromady $2b + 1$ dětí, můžeme rozdělit na tři vrcholy, kde každý má $2b/3$ dětí. (Jednoho sourozence máme vždy, neboť $a \geq 2$.)
Ale takto můžeme spojit tři malé vrcholy do dvou – pro každý vrchol, který má po odebrání $a - 1$ synů se nejprve pokusíme ukrást jednoho syna svému sourozenci. Pokud oba sousední sourozenci mají přesně $2b/3$ dětí, pak je spojíme do dvou vrcholů. Tady je také potřeba dát si speciální pozor na případ, kdy jsme odebrali syna nejlevějšímu vrcholu - pak lze například ukrást jednoho syna svému sousedovi, a přesunout se k němu, čímž převádíme na případ s oběma sousedními sourozenci.