3. CVIČENÍ Z DATOVEK 1, ZS26/27

Splay stromy

1. *Naivní rotace hledaného prvku do kořene.* Upravíme BVS (bez vyvažování) tak, že po vyhledání klíče x uzel s klíčem x *vyrotujeme* do kořene, tedy opakujeme následující: dokud má vrchol x rodiče p, provedeme jednoduchou rotaci mezi x a p. (Pro jednoduchost žádné klíče nevkládáme ani nemažeme.) Jinými slovy: vezmeme splay strom a místo dvourotací (zig-zig a zig-zag) děláme pouze jednoduché rotace.

a) *Přístupy v rostoucím pořadí.* Začněte s libovolným binárním vyhledávacím stromem obsahujícím právě klíče $1, 2, \ldots, n$, kde $n \geq 2$. Postupně proveďte přístupy ke klíčům $1, 2, \ldots, n$, vždy v aktuálním stromu vzniklém po předchozím přístupu. Jak bude vypadat strom po celé posloupnosti? Závisí jeho výsledný tvar na počátečním stromu?

b) *Opakování rostoucí posloupnosti.* Začněte se stromem, který vznikl na konci předchozí úlohy, a znovu proveďte přístupy $1, 2, \ldots, n$. Určete celkový počet rotací a časovou složitost pro tuto druhou rostoucí posloupnost.

c) *Splay to the rescue!* Zkuste předchozí dva body na správném splay stromu se správnými dvourotacemi. Co dostanete po první posloupnosti? Jak se bude lišit druhý průchod od naivních rotací?

2. *Splay vs vyvažování.* Srovnejte splay stromy s dalšími BVS. Jaké jsou výhody a nevýhody splay stromů?

3. *Splay vs staticky optimální strom.* Může být splay strom asymptoticky rychlejší než staticky optimální strom či AVL strom na nějaké posloupnosti přístupů?

4. *Splay potenciál.* Ujistěte se, že chápete, jak je definovaný potenciál ve splay stromech a že rozumíte hlavním myšlenkám analýzy amortizované složitosti splaye.

a) Jak je definován potenciál splay stromu?

b) Jaký je potenciál perfektně vyváženého stromu? (Stačí nám rozumný horní a dolní odhad, pro jednoduchost předpokládejte $n = 2^k - 1$.)

c) Jaký je potenciál cesty? (Opět stačí rozumné odhady.)

d) Jaká je amortizovaná cena rotace (dvojité a jednoduché)?

e) Jaká je amortizovaná cena celého splaye a jak plyne z amortizovaných cen rotace?

f) Jaké je *reálná* cena celého splaye a v jakých jednotkách se jí hodí počítat?

5. *Bonus: Potenciál pro následníka.* Už jsme si dokázali, že použití n operací následníka na libovolném BVS má složitost $O(n)$ (když začneme ve vrcholu s nejmenším klíčem). Jak to dokázat za pomoci potenciálu, tedy jaký potenciál zvolit?

6. *Bonus: Splay spojový seznam* Uvažme následující problém: máme spojový seznam, a v něm máme prvky $x_1, \ldots, x_n$. Dále pro každý prvek x_i víme, že pravděpodobnost dotazu na tento prvek je p_i, kde $\sum_{i=1}^{n} p_i = 1$. Cena dotazu na prvek x_i je jeho pořadí ve spojovém seznamu. Mějme prvky x_i označené tak, že $p_1 \geq p_2 \geq \ldots \geq p_n \geq 0$.

a) Ukažte, že mezi všemi *fixními* spojovými seznamy je námi vytvořený spojový seznam v pořadí $x_1, x_2, \ldots, x_n$ dle pravděpodobností ve střední hodnotě optimální. Tuto hodnotu dále značíme jako $\mathbb{E}[T_{\text{opt}}]$.

b) Uvažme dále tzv. *move-to-front* spojový seznam, kde v každém kroku prvek ve spojovém seznamu vyhledáme, a potom ho umístíme na začátek. (Na začátku je pořadí prvků libovolné.) Ukážeme, že v tomto případě bude střední hodnota ceny maximálně dvojnásobná, a provedeme to následovně. Prvně budeme předpokládat, že na každý prvek jsme se už alespoň jednou podívali, takže pořadí prvků ve spojáku nezávisí na původním pořadí, a budeme zkoumat pořadí prvků v nějaký fixní moment v čase. Z prvního kroku máme, že $\mathbb{E}[T_{\text{opt}}] = \sum_{i=1}^{n} i \cdot p_i$, a dále chceme spočítat $\mathbb{E}[T_{\text{mtf}}] = \sum_{i=1}^{n} \mathbb{E}[R_i] \cdot p_i$. kde R_i je náhodná veličina popisující pořadí prvku x_i v náš konkrétní moment. Konkrétně $\mathbb{E}[R_i] = 1 + \sum_{j=1, j\neq i}^{n} 1 \cdot \Pr[x_j \text{ je před } x_i]$ z linearity střední hodnoty. Vaším úkolem je nyní jen určit pravděpodobnost $\Pr[x_j \text{ je před } x_i]$, a dále odhadnout $\mathbb{E}[T_{\text{mtf}}] \leq 2 \cdot \mathbb{E}[T_{\text{opt}}]$.