3. CVIČENÍ Z DATOVÝCH STRUKTUR 1, ZS24

BVS vyvažované líně nebo naopak velmi proaktivně (pomocí splay operací)

1. *Konstrukce optimálního statického BVS.* Mějme množinu klíčů $k_1 < k_2 < \cdots < k_n$, ze kterých chceme postavit statický BVS. Nechť známe dopředu četnosti přístupů $w_1, \ldots, w_n$ k jednotlivým klíčům (které lze také interpretovat jako pravděpodobnosti přístupů). Postavte *optimální BVS*, tedy BVS minimalizující

$$\sum_{i=1}^{n} h(i) \cdot w_i$$

kde $h(i)$ je hloubka klíče k_i v postaveném BVS (kořen má hloubku 0). Jinými slovy provedení w_i přístupů ke klíči k_i pro každé i (beze změny stromu) bude trvat co nejmenší dobu.

Hint: dynamické programování.

2. *Líně vyvažované stromy, neboli BB[α].* Připomeňte si, jak fungují líně vyvažované BVS a jakým potenciálem se analyzují.

a) Jak dlouho může trvat provedení k operací provedených na libovolném BB[α] stromu s n vrcholy?
b) Proč je v definici potenciálu výjimka pro rozdíl 1, tedy co by se pokazilo, kdybych ji neudělali?

3. *Hloubka splay stromů nemusí být vždy logaritmická.* Navrhněte posloupnost operací, která vytvoří splay strom s lineární hloubkou. Můžete nejdřív začít s prázdným stromem a pak s libovolným BVS.

4. *Splay naivně.* Co se pokazí, pokud splay stromy implementujeme naivně, tedy jen pomocí jednoduchých rotací jedné hrany?

Hint: zkuste cestu. Jak se chovají dvojité rotace na cestě?

5. *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) Jaká je amortizovaná cena rotace (dvojité a jednoduché)?
c) Jaká je amortizovaná cena celého splaye a jak plyne z amortizovaných cen rotace?
d) Jaké je *reálná* cena celého splaye a v jakých jednotkách se jí hodí počítat?
e) Jaký je potenciál perfektně vyváženého stromu? (Stačí nám rozumný horní a dolní odhad, pro jednoduchost předpokládejte $n = 2^k - 1$.)
f) Jaký je potenciál cesty? (Opět stačí rozumné odhady.)

6. *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?

7. *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 $\mathbb{E}[T_{\text{mtf}}] \leq 2 \cdot \mathbb{E}[T_{\text{opt}}]$.