

3. cvičení

Datové struktury I, 16. 10. 2025

<https://iuuk.mff.cuni.cz/~chmel/2526/ds1/>

Úloha 1 (Splay stromy mají potenciál)

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

- Jak je definován potenciál splay stromu?
- Jaký je potenciál perfektně vyváženého stromu? (Stačí nám rozumný horní a dolní odhad, pro jednoduchost předpokládejte $n = 2^k - 1$.)
- Jaký je potenciál cesty? (Opět stačí rozumné odhady.)
- Jaká je amortizovaná cena rotace (dvojitě a jednoduché)?
- Jaká je amortizovaná cena celého splaye a jak plyne z amortizovaných cen rotace?
- Jaká je *reálná* cena celého splaye (a v jakých jednotkách ji vlastně počítáme)?

Úloha 2 (Semisplay)

V krocích LL a PP (neboli zig-zig) můžeme být línější a místo provedení dvojrotace provést jen jednoduchou rotaci vrchní hrany. Konkrétně v případě LL mějme tři vrcholy x, y a z takové, že x je levým synem y a y je levým synem z . Pak při splayování x provedeme jednoduchou rotaci hrany zy a dále pokračujeme ve splayování vrcholu y (místo x); původně splayovaný prvek se tedy *nemusí dostat* do kořene. Kroky LP a PL (zig-zag) nebo závěrečná jednoduchá rotace fungují stejně. Zanalyzujte amortizovanou cenu operace semisplay.

Úloha 3 (Ultralíné splay stromy)

Motivováni úspěšnou leností, která vedla k operaci semisplay, chceme na stromě provádět ještě méně změn. Co kdybychom při splayování prvku x rotovali jen každou druhou hranu na cestě z x do kořene a postupně měnili splayovaný prvek?

Úloha 4 (Rozděl a spojuj)

Pro splay strom T a hodnotu k navrhnete operaci SPLIT, která strom T rozdělí na dva stromy T', T'' , přičemž ve stromě T' jsou všechny hodnoty menší nebo rovny k a ve stromě T'' jsou všechny hodnoty větší než k . Pokuste se zachovat amortizovanou složitost.

Dále pro splay stromy T', T'' , kde všechny hodnoty v T' jsou menší než všechny hodnoty v T'' , navrhnete operaci JOIN(T', T''), která sloučí stromy do jednoho, opět při zachování amortizované složitosti.

Bonusové úlohy

Úloha 5 (Subset theorem)

Ukážeme následující tvrzení: mějme splay strom na vrcholech $[n]$ a podmnožinu $S \subseteq [n]$, $s := |S|$. Pak m dotazů, které se ptají pouze na prvky S , má časovou složitost $\mathcal{O}(n \cdot s + m \log s)$.

Jako bonus jej můžeme zlepšit na $\mathcal{O}(n \cdot \min(s, \log n) + m \log s)$.

Neformálně důkaz postupuje následovně: rádi bychom si představili, že prvky S nám ve stromě tvoří malý podstrom (ideálně ve velikosti lineární v s), ve kterém se děje skoro vše důležité. Tak takový podstrom budeme sledovat s tím, že se tedy jedná o nejmenší podstrom, který obsahuje všechny prvky, na které jsme se dosud ptali. Ukážeme, že přidání jednoho nového vyhledaného prvku nám do podstromu přidá maximálně jeden další prvek, takže celý podstrom nebude příliš velký. Když už prvek v tom podstromě je, tak cenu všech dotazů můžeme počítat jenom v tom malém podstromě, ale když v něm ještě není, tak musíme cenu rozdělit na cenu mimo podstrom, a cenu v podstromě. Mimo podstrom můžeme cenu odhadnout triviálně. V podstromě pak můžeme použít větu o amortizované složitosti, a díky malé velikosti podstromu $\mathcal{O}(s)$ dostáváme celkový výsledek.

Předpokládejme, že máme posloupnost dotazů $q_1, \dots, q_m$. Dále si označíme jako $T_0, \dots, T_m$ stavy splay stromu, kde T_i je stav po i -té operaci, a jako $S_i := \{q_j : j \leq i\}$, $s_i := |S_i|$ si označíme množinu prvků, které byly dotazovány do i -té operace, a mohutnost této množiny. Budeme postupovat následovně:

- Ukážeme, že když se podíváme na nejmenší podstrom T_i , který obsahuje všechny prvky S_i , tak tento podstrom obsahuje nejvýše $2s_i - 1$ vrcholů. Tento podstrom si dále označíme jako D_i . (D_0 je prázdný strom.)

- b) Ukážeme, že to, co se děje při rotacích mezi T_i a T_{i+1} v D_i umíme simulovat v malém stromě, který odpovídá D_i .
- c) Tím pádem můžeme cenu každého dotazu rozdělit na dvě části: cena na „zvednutí“ do D_i a cena uvnitř D_i . Triviálně odhadněte cenu na zvednutí s prvků jako $\mathcal{O}(n \cdot s)$.
- d) Odhadneme za pomoci nevážené analýzy cenu všech operací uvnitř $D = (D_0, \dots, D_m)$ (z prvního bodu máme odhad na velikost D_m) a uvědomíme si, že to stačí.

Bonus: Využijme věty ze cvičení a uvědomme si, že pro velké s ji také můžeme použít se stejným výsledkem.

Bonus 2: Jak to vypadá s částí a) pro naivní splay?