

3. CVIČENÍ Z ADS 2, ČTVRTEK 15:40

Vyhledávání více jehel v kupce sena

1. *Aho-Corasicková*: Sestrojte vyhledávací automat pro slova dar, radar, adam, a.
2. *Příliš mnoho výskytů*: Nalezněte příklad jehel a sena, v němž je asymptoticky více než lineární počet výskytů. Přesněji řečeno ukažte, že pro každé n existuje vstup, v němž je součet délek jehel a sena $\Theta(n)$ a počet výskytů není $\mathcal{O}(n)$.
3. *Naivní skákání*: Uvažujme zjednodušený algoritmus AC, který nepoužívá zkratkové hrany a vždy projde po zpětných hranách až do kořene. Ukažte vhodnými příklady vstupů, že tento algoritmus je asymptoticky pomalejší.
4. *Frekvence výskytů*: Popište algoritmus, který v lineárním čase pro každou jehlu spočítá, kolikrát se v seně vyskytuje. Časová složitost by neměla záviset na počtu výskytů – ten, jak už víme, může být superlineární.
5. *Předpočítané množiny*: Jednoduchý způsob, jak si poradit s hlášením výskytů, je předpočítat si pro každý stav s množinu $\text{out}(s)$ slov (jehel) k ohlášení. Dokažte, že tyto množiny není možné sestavit v lineárním čase s velikostí slovníku, protože součet jejich velikostí může být pro některé vstupy superlineární. (Poznámka: Množiny $\text{out}(s)$ obsahují jen *indexy* jehel ve slovníku, nikoliv celá slova.).
6. *Dynamické vyhledávání*: Navrhněte datovou strukturu pro dynamické vyhledávání v textu. Jehla je pevná, v seně lze průběžně měnit jednotlivé znaky a struktura odpovídá, zda se v seně právě vyskytuje jehla.

7. *Dvoudimenzionální hledání:* Jak rozhodneme jestli matice A o rozměrech $n \times n$ obsahuje jako souvislou podmatici matici B o rozměrech $m \times m$. (Hint: pro začátek uvažujte, že sloupce (nebo řádky) matice B jsou různé.)

Bonus^d: uvažujte d -dimenzionální matice.

8. *Nejmenší slovo bez výskytu:* Na vstupu dostanete množinu slov ι_1, \dots, ι_n nad anglickou abecedou a přirozené číslo l . Najděte lexikograficky nejmenší slovo délky l , které neobsahuje ι_i pro každé i .