

Drugi projekat - FAQ

Šta tačno predstavlja zadatak 2 - simulacija pozivanja iz fajla?

Zadatak 2 je osmišljen kao automatizacija zadatka 1. Umesto unošenja brojeva, podaci o pozivu se učitavaju iz fajla. Fajl koji se koristi ne mora biti veličine calls.txt, može i neki značajno manji. Sumarni podaci se odnose na osnovne podatke o pozivu.

Da li se očekuje da se brojevi telefona sastoje isključivo od cifara?

Ne nužno. Često, kada zapisujemo brojeve telefona dodajemo razmake i crtice kako bismo broj lakše zapisali/zapamtili. Ukoliko se u sklopu broja nađu crtice ili razmaci trebalo bi da se ignorišu. Ukoliko se unutar broja nađu slova ili neki drugi znaci koji nemaju značenje, broj se smatra neispravnim pa bi taj poziv trebalo ignorisati.

Kako mogu da osvojim dodatne poene?

Dodatni zadaci nisu definisani unapred nego se očekuje od studenata da ih predlože. Asistent može da odobri zadatak i definiše koliko dodatnih poena nosi. Takav zadatak će biti dopisan u tekst projektnog zadatka.

Da li postoji vremensko ograničenje kod učitavanja velikih fajlova npr. calls.txt?

Poželjno je da se veliki fajlovi pretprocesiraju i da se pre same odbrane formiraju potrebne strukture podataka. Te strukture treba da se uskladište u memoriji i učitaju prilikom pokretanja programa. Vreme od pokretanja do izvođenja prve operacije u aplikaciji ne bi trebalo da traje duže od nekoliko sekundi. Neprihvatljivo je da se pri svakoj pretrazi podaci ponovo učitavaju, strukture formiraju i slično.

U kojim zadacima je obavezno koristiti strukturu podataka trie?

Trie je prefiksno stablo tako da će značajno ubrzati pretragu gde god se vrši prefiksna pretraga ili pretraga po celokupnoj vrednosti (primarno zadaci 5 i 8, ali i 1, 2, 3, 4). Gde god postoji problem pronalaženja broja telefona u grafu, potrebno je taj broj (čvor) efikasno pronaći.

Zadatak 5 - da li postoje 3 različite pretrage ili je u pitanju jedna pretraga?

Postoje 3 nezavisne pretrage a na početku se bira kriterijum pretraživanja.

Da li blokirane brojeve treba ignorisati čak i prilikom učitavanja iz fajla calls.txt?

Ne. Ako je broj telefona trenutno blokiran, ne može da učestvuje u pozivima ali ne bi trebalo da se potpuno ukloni iz evidencije - grafa popularnosti.

U kojim zadacima koristimo graf popularnosti?

U svim zadacima gde je potrebno utvrditi veze - pozive između brojeva telefona, dakle u skoro svim.

Da li opcija 'did you mean' zahteva da bar prvih par cifara brojeva budu iste?

Ne, ne moraju. U pythonu postoji nekoliko dostupnih modula koje se bave računanjem sličnosti (similarity). Implementirajte sopstveni algoritam ili iskoristite postojeće odgovarajuće implementacije.

Da li odmah prilikom implementacije grafa popularnosti računam i popularnost svakog broja, što onda znači da i taj podatak treba da smestim u svoj graf?

Za popularnost broja u grafu je u redu i opcija da se popularnost broja izračuna na osnovu grafa u trenutku kada je ta informacija potrebna, i da se skladišti u čvoru na neki način.

Gde da čuvam informacije vezane za poziv koje utiču na popularnost?

Informacije vezane za poziv koje utiču na popularnost nije neophodno čuvati ako ih odmah nakon poziva iskoristite da ažurirate veze u grafu kao i podatke vezane za te veze. Ako ipak želite, možete neke neophodne informacije vezivati za Edges.

Da li mogu da pravim 3 posebna trie-a za imena, prezimena i brojeve ili sve podatke treba da čuvam u jednom?

Ispravno je formirati 3 odvojena trie-a. Dozvoljeno je formirati i jedan trie uz osmišljanje načina za razlikovanje imena i prezimena.