

Drugi veliki domaći zadatak

Napomene:

- 1) Odbrana: sreda, 19.12.2018. godine u 16:30, laboratorija 60.
- 2) Bira se lakši ili teži problem.
- 3) Rešenje jednog od ova dva problema treba doneti na odbranu, u vidu fajla dz2.py.
- 4) U zaglavlju fajla, pod komentarom, treba navesti ime, prezime i broj indeksa.
- 5) Rešenja treba da budu otporna na greške, odnosno program treba da detektuje i prijavi pogrešno unete vrednosti (van predviđenog opsega ili pogrešnog tipa, nepostojeći fajl).
- 6) Potrebno je osmisliti test primere i doneti ih na odbranu.

Lakši problem

Napisati funkciju na jeziku Python koja modifikuje fajl čiji je naziv zadat argumentom. Modifikacija podrazumeva sledeće izmene fajla:

- Svaku liniju teksta u fajlu koja je duža od 80 karaktera treba "prelomiti", tj. pretvoriti je u više linija koje sadrže ne više od 80 karaktera, ali tako da se tekst ne prelama u sred reči.
- Izbaciti pojavljivanje viška blanko znakova (više razmaka zameniti jednim, više praznih redova zameniti jednim i slično).
- Pre interpunkcijskih simbola (tačka, zarez, uzvičnik, upitnik, dvotačka, tačka-zarez, ...) ne sme da se nalazi blanko znak (razmak), a nakon simbola se mora naći razmak. Izuzetak su datumi i vremena (18:00 treba da ostane u tom formatu, 21.12.2018. takođe).
- Pojavljivanje crtice ("-") u tekstu rešiti na sledeći način: ako se neposredno pre ili posle crtice nalazi razmak, onda sa svake strane treba da se nađe po jedan razmak; ako ne postoji razmak oko crtice, ostaviti tako kako jeste (pretpostavljamo da su to polusloženice).
- Svaka rečenica mora početi velikim slovom.

Modifikovan fajl treba sačuvati pod drugim nazivom, tako da se na originalan naziv fajla doda sufiks "-mod". Kratak primer ulaznog fajla i modifikovanog fajla se nalazi u prilogu, kao i veliki tekst koji možete modifikovati (fajl "big.txt").

Napisati program koji od korisnika traži unos imena fajla koji treba modifikovati i poziva funkciju za obradu, sve dok korisnik ne unese tekst "kraj".

Teži problem

Posmatra se šifrovana komunikacija između dve učesnika. Jedan učesnik drugom šalje fajlove koji sadrže šifrovani tekst. Tekst se šifrira korišćenjem Cezarove šifre, sa nepoznatim pomakom, koji je prilikom svakog slanja drugačiji. Cezarova šifra je šifra supstitucije, tj. svaki karakter se menja karakterom koji je nalazi za p pozicija nakon njega u abecedi, gde je p pomeraj. Za potrebe ove komunikacije, posmatra se proširena abeceda, koja se sastoji samo od velikih slova, cifara i specijalnih znakova. Abeceda koja se koristi data je u sledećem stringu: "ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789" '- , . : ; ? ! () [] _ " (primetiti da postoji znak razmaka na kraju). Tako na primer, ako se šalje rečenica „PROSTO K’O PASULJ“, sa pomerajem 10, dobija se poruka “Z1Y23YJU[YJZK24VT”.

Zadatak je dešifrovati dobijene fajlove. Primena „brute-force“ rešenja i provera da li za svaki pomeraj tekst ima smisla nije baš dobro rešenje, ali sa sigurnošću znamo da fajlovi poslani od strane iste osobe imaju određeni zajednički sadržaj, tj. postoji deo teksta koji će se pojavljivati u svakom od poslanih fajlova. Takođe je poznato da se zajednički sadržaj nalazi na početku ili na kraju fajla (kao *header* ili *footer*).

Dešifrovan fajl treba da se zove isto kao šifrovani, sa sufiksom "_dc" (skraćeno od *decrypted*).

Dozvoljena je i određena interakcija sa programom, u smislu da se može funkciji koja dešifruje prosledivati koji deo teksta treba da proba da dešifruje, da bi nakon toga mogla da dešifruje ceo tekst.

U prilogu VDZ2_test.zip nalaze se šifrovani fajlovi i jedan dešifrovan (tj. „originalni“ fajl).