

Interpretacija programa

7. rujna 2018.

Ovo je "open book" kolokvij. Dozvoljeno je korištenje bilo kakvih materijala — bilješke s vježbi/predavanja, vlastiti USBovi s riješenim zadacima, Python help, tutoriali, postovi na online forumima,... — **nastalih prije** kolokvija (npr. dozvoljeno je na *StackOverflowu* naći rješenje nekog zadatka, ali nije dozvoljeno tamo postaviti pitanje kako se rješava neki zadatak). Također, nije dozvoljena komunikacija (razgovor, *chat*, razmjena papira, USBova, ili bilo kakvih materijala) **među** studentima.

U Github repozitoriju kolegija nalazi se zadnja verzija biblioteke `pj.py`. Rješenje zadatka pišite **isključivo** u lokalnu kopiju datoteke `pj_kemija.py`, te na kraju tu datoteku pošaljite emailom na veky@math.hr, sa subjectom "IP K2". Odgovore na pitanja koja se nalaze u zadatku pišite u komentare unutar koda.

Datoteka `pj_kemija.py` treba se moći izvršiti u Pythonu bez grešaka. Ako imate neki kod koji po Vašem mišljenju pokazuje ideju rješenja, ali iz nekog razloga ne radi, napišite ga u komentar. Korisno je u datoteku uključiti i testove koji pokazuju kako kod radi (ili ne radi).

Maksimalno vrijeme rješavanja je 180 minuta. Na kolokviju je moguće osvojiti najviše 30 bodova.

Formule za kemijske spojeve sastoje se od simbola za atome, brojeva i zagrada (obljih i uglatih), te eventualno varijable n . Recimo, natrijev trikarbonat kobaltat ima formulu $\text{Na}_3[\text{Co}(\text{CO}_3)_3]$, dok zasićeni ugljikovodici imaju opći oblik formule $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$. Sve subskripte pišemo normalno kao brojeve, recimo $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$. (Za kemičare: ne promatramo izotope, ione niti hidrate. Ako ne znate što je to, još bolje.)

Broj nakon simbola atoma znači toliki broj pojavljivanja atoma u molekuli. Broj ili n nakon zatvorene zagrade znači toliki broj pojavljivanja skupine atoma koja je u zagradi. Broj mora biti prirodni (pozitivan). Ako broja nema, podrazumijeva se 1. Recimo, $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$ ima $n+2$ ugljikovih atoma (C), i $2n+6$ vodikovih atoma (H). Simboli atoma imaju ili jedno veliko slovo (npr. O), ili jedno veliko i jedno malo slovo (npr. Na).

[5b] Napišite leksički analizator za jezik kemijskih formula. Praznine u formulama nisu dozvoljene. [5b] Potrebne tipove tokena odredite sami. Pažnja: Varijabla n se može pojaviti samo neposredno nakon zatvorene zagrade. Sn je kositar, $(S)_n$ je n atoma sumpora.

Napišite [5b] beskontekstnu gramatiku i [10b] sintaksni analizator za jezik kemijskih formula. Apstraktna sintaksna stabla odredite sami, ali moraju sadržavati dovoljno informacija da iz njih bude moguće rekonstruirati formulu.

[5b] Napišite semantički analizator u obliku kalkulatora molne mase spoja. Realizirajte ga kao metodu na apstraktnim sintaksnim stablima, koja prima relativne atomske mase svih atoma koji se u spoju pojavljuju, i eventualno vrijednost za n . Primjerice, spoj (' $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$ ').Mr(C=12.01,H=1.008,n=2) će dati molnu masu butana, koja iznosi $4*12.01+10*1.008=58.12$.

Na <https://www.webqc.org/mmcaltc.php> možete provjeriti svoja rješenja (samo morate uvrstiti neku konkretnu vrijednost za n).