

Skriptni jezici: zadatak za izlazni test

Napomene

- Skriptu imenovati **zadatak1.py**
- Argumenti: URL stranice, ime log datoteke
- Izlaz: generirani podaci ispisuju se na stdout
- Isprobanu skriptu predati kroz [Fferko](#).
- Obavezno zaključati upload!
- Prilikom rješavanja izlaznog testa nije dozvoljeno koristiti Internet ni komunikaciju s kolegama!
- Pri pisanju kôda voditi računa o preglednosti.
- Skripta mora biti sintaksno ispravna i mora se moći izvesti u Pythonu 3.

Zadatak

Napisati Python skriptu koja će u log-datoteci web poslužitelja analizirati broj pristupa zadanoj stranici. Kao argumenti se navode ime stranice koja se analizira i ime log datoteke. Skripta treba ispisati broj pristupa s pojedinih IP adresa i to samo za one adrese s kojih je bilo barem 2 pristupa zadanoj stranici, sortirano silazno prema broju pristupa.

Skripta treba provjeriti broj argumenata, te uspješnost otvaranja navedene log datoteke. U slučaju neispravnog pozivanja ispisati uputu o načinu korištenja i izaći iz skripte.

Ispis oblikovati prema primjeru u nastavku, a za testiranje se može koristiti log-datoteka iz laboratorijske vježbe ([localhost_access_log.2008-02-24.txt](#)).

Dodatni naputci:

- Za dohvat argumenata naredbenog retka može se koristiti standardni modul sys. Argumenti su pohranjeni u listi sys.argv.
- Silazno sortiranje liste n-torki po željenom članu n-torke može se postići koristeći sljedeći "recept":
sortirano = sorted(entorke, key=lambda t: t[1], reverse=True)

Poziv skripte zadatak1.py:

```
$ python3 zadatak1.py "/zadaci-faq.html" TestPrimjeri/localhost_access_log.2008-02-24.txt
```

```
-----  
Broj pristupa stranici: /zadaci-faq.html  
IP adresa pristupatelja : Broj pristupa  
-----
```

```
89.164.250.5 : 15  
193.198.16.211 : 13  
89.164.46.53 : 9  
89.164.7.174 : 8  
89.164.228.250 : 3  
78.1.106.36 : 2  
78.0.25.238 : 2  
89.172.2.247 : 2  
90.157.248.22 : 2  
78.1.111.73 : 2  
83.131.54.227 : 2  
90.157.251.180 : 2  
89.172.56.25 : 2  
83.131.40.124 : 2  
89.172.20.163 : 2  
89.172.53.0 : 2  
-----
```

"Skriptni jezici", FER 2015/16

Copyright(c) [Zavod za elektroniku, mikroelektroniku, računalne i inteligentne sustave](#)
[Fakultet elektrotehnike i računarstva](#)
[Sveučilište u Zagrebu](#)

Zoran Kalafatić
Zadnja promjena: 10.06.2016.

Skriptni jezici: zadatak za izlazni test

Napomene

- Skriptu imenovati **zadatak3.py**
- Argumenti: imena HTML datoteka
- Izlaz: generirani podaci ispisuju se na `stdout`
- Isprobano skriptu predati kroz [Fferko](#).
- Obavezno zaključati upload!
- Prilikom rješavanja izlaznog testa nije dozvoljeno koristiti Internet ni komunikaciju s kolegama!
- Pri pisanju kôda voditi računa o preglednosti.
- Skripta mora biti sintaksno ispravna i mora se moći izvesti u Pythonu 3.

Zadatak

Napisati Python skriptu koja će kao argument naredbenog retka prihvatiti proizvoljan broj imena HTML datoteka. Skripta mora proći po svim zadanim datotekama, te na standardni izlaz ispisati koliko ukupno pojedinih HTML tagova ima unutar **svih** datoteka. Sva imena tagova potrebno je normalizirati u mala slova. Ispis sortirati prema imenima tagova i oblikovati prema primjeru.

Neke dodatne napomene za ostvarivanje zadatka:

- Pripaziti da se broje i elementi koji imaju atribut!
- Za dohvat argumenata naredbenog retka može se koristiti standardni modul `sys`. Argumenti su pohranjeni u listi `sys.argv`.

Primjer izvršavanja skripte:

```
$ python3 zadatak3.py stranica1.html index.html
```

```
Statistika tagova:
Tag          Broj ponavljanja
=====
a              8
body           2
br            10
code           8
font          12
h1             2
h2             4
head           2
hr             6
html           2
i              2
li            12
meta           2
p             16
pre            2
table          6
td             8
title          2
tr             6
ul             4
```


Zadatak

Uobičajen zapis podataka o znanstvenoj literaturi je `bibtex`. U prilogu je jedna takva datoteka ([knjige.bib](#)), s pojednostavljenim formatom. Koristi se samo jedna vrsta zapisa (`@book`), te samo 4 vrste polja (`author`, `publisher`, `title` i `year`).

Potrebno je napisati Python skriptu koja će na temelju `HTML` datoteke u kojoj se nalaze zapisi o knjigama u obliku (oznaka je šifra koja se navodi na početku odgovarajućeg `bib` zapisa):

```
<UL>
  <LI> author, <I>title</I>, publisher, year (oznaka) </LI>
  <LI> sljedeća knjiga </LI>
  ...
</UL>
```

generirati (i ispisati na standardni izlaz) odgovarajuću `bib` datoteku. Ime ulazne `HTML` datoteke navodi se kao argument naredbenog retka. Primjer ulazne datoteke je: [knjige.html](#) Skripta treba provjeriti broj argumenata, te uspješnost otvaranja navedene datoteke s podacima. U slučaju neispravnog pozivanja ispisati uputu o načinu korištenja i izaći iz skripte.

Ispis oblikovati prema primjeru u nastavku.

Dodatni naputci:

- Za dohvat argumenta naredbenog retka može se koristiti standardni modul `sys`. Argumenti su pohranjeni u listi `sys.argv`.

Primjer pozivanja skripte:

```
$ python3 zadatak4.py TestPrimjeri/knjige.html
```

```
@book{DudaHart,
  author = "R. O. Duda and P. E. Hart and D. G. Stork",
  title = "Pattern Classification",
  publisher = "Wiley",
  year = "2000"
}
```

```
@book{shapiro01,
  author = "Linda G. Shapiro and George C. Stockman",
  title = "Computer Vision",
  publisher = "Prentice Hall",
  year = "2001"
}
```

```
@book{Forsyth02,
  author = "David A. Forsyth and Jean Ponce",
  title = "Computer Vision: A Modern Approach",
  publisher = "Prentice Hall",
  year = "2002"
}
```