Programmazione in Python

Esercizi

Dario Pescini - Mirko Cesarini $^{ m 1}$

Università degli Studi di Milano-Bicocca
Dipartimento di Statistica e Metodi Quantitativi

dario.pescini@unimib.it

Indovina la parola

Si chiede di implementare un programma che, scelta una parola a caso tra quelle presenti nel file listaParoleIta.txt, chieda all'utente di indovinarla.

Il programma dovrà, inoltre, fornire come aiuto il conteggio, per ciascuno dei caratteri presenti nella parola proposta dall'utente, delle sue occorrenze nella parola da indovinare.

```
ossequiosi

la parola è composta da 12 caratteri
che parola è? (-1 per terminare) computer
aiuto:
c e' presente 0 volte
e e' presente 1 volte
m e' presente 0 volte
o e' presente 2 volte
p e' presente 0 volte
r e' presente 0 volte
u e' presente 1 volte
u e' presente 0 volte
t e' presente 0 volte
```

file listaParoleIta.txt

```
a
abate
abati
abbagli
abbaglia
abbagliai
abbagliamo
abbagliano
abbagliare
abbagliare
abbagliate
abbagliate
abbagliate
```

zuffoli zuffoliamo zuffolino zuffolo zuppa zuppe zuppiera zuzzurellone

Soluzione

```
fileName = 'listaParoleIta.txt'
# caricamento lista parole possibili target
fileParole = open(fileName, 'r')
listaParole = fileParole.readlines()
fileParole.close()
lunghezzaLista = len(listaParole)
# sorteggio della parola segreta
iParola = random.randint(0, lunghezzaLista - 1)
parolaDaIndovinare = listaParole[iParola]
# print parolaDaIndovinare
print 'la parola è composta da ', len(parolaDaIndovinare), 'caratteri'
```

```
while termina == False and indovinato == False:
    parolaTentativo = raw_input('che parola è? (-1 per terminare) ')
    if parolaTentativo == '-1':
        if parolaTentativo == parolaDaIndovinare:
            print 'hai indovinato'
            conteggio = {}
            for carattere in parolaTentativo:
                 if carattere not in conteggio.keys():
                     conteggio[carattere] =
    parolaDaIndovinare.count(carattere)
            print "aiuto:"
            for el in conteggio.keys():
                 print el, "e' presente ", conteggio[el], 'volte'
```

Codice Fiscale

Si chiede di implementare in python l'algoritmo usato per la generazione del codice fiscale.

I primi quindici caratteri sono indicativi dei dati anagrafici di ciascun soggetto secondo l'ordine seguente:

- tre caratteri alfabetici per il cognome;
- tre caratteri alfabetici per il nome;
- due caratteri numerici per l'anno di nascita;
- un carattere alfabetico per il mese di nascita;
- due caratteri numerici per il giorno di nascita ed il sesso;
- quattro caratteri, di cui uno alfabetico e tre numerici per il comune italiano o per lo Stato estero di nascita.

Il sedicesimo carattere, alfabetico, ha funzione di controllo.

Generatore codice fiscale: caratteri indicativi del cognome

- I cognomi che risultano composti da più parti o comunque separati od interrotti, vengono considerati come se fossero scritti secondo un'unica ed ininterrotta successione di caratteri.
- Per i soggetti coniugati di sesso femminile si prende in considerazione soltanto il cognome da nubile.
- Se il cognome contiene tre o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il cognome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il cognome contiene una consonante e due vocali, si rilevano, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il cognome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il cognome e' costituito da due sole vocali, esse si rilevano, nell'ordine, e si assume come terzo carattere la lettera x (ics).

Generatore codice fiscale: caratteri indicativi del nome

- I nomi doppi, multipli o comunque composti, vengono considerati come scritti per esteso in ogni loro parte e secondo un'unica ed ininterrotta successione di caratteri.
- Se il nome contiene quattro o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la terza e la quarta consonante.
- Se il nome contiene tre consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il nome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il nome contiene una consonante e due vocali, i tre caratteri da rilevare sono, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il nome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il nome e' costituito da due sole vocali, esse si rilevano nell'ordine, e si assume come terzo carattere la lettera x (ics).

Generatore codice fiscale: Data, sesso e luogo di nascita

- I due caratteri numerici indicativi dell'anno di nascita sono, nell'ordine, la cifra delle decine e la cifra delle unità dell'anno stesso.
- Il carattere alfabetico corrispondente al mese di nascita e' quello stabilito per ciascun mese nella seguente tabella:

Gennaio = A	Maggio = E	Settembre = P
Febbraio = B	Giugno = H	Ottobre = R
Marzo = C	Luglio = L	Novembre = S
Aprile = D	Agosto = M	Dicembre = T

Generatore codice fiscale: Data, sesso e luogo di nascita

- I due caratteri numerici indicativi del giorno di nascita e del sesso vengono determinati nel modo seguente:
 - Per i soggetti maschili il giorno di nascita figura invariato, con i numeri da uno a trentuno, facendo precedere dalla cifra zero i giorni del mese dall'uno al nove.
 - Per i soggetti femminili il giorno di nascita viene aumentato di quaranta unità, per cui esso figura con i numeri da quarantuno a settantuno.
- I quatto caratteri alfanumerici indicativi del comune italiano o dello Stato estero di nascita, sono costituiti da un carattere alfabetico seguito da tre caratteri numerici, secondo la codifica stabilita dall'Agenzia del Territorio e contenuta nel file CodiceComuniItalia.csv

Generatore codice fiscale: Carattere alfabetico di controllo

Il sedicesimo carattere ha funzione di controllo dell'esatta trascrizione dei primi quindici caratteri e viene determinato in questo modo:

 ciascuno degli anzidetti quindici caratteri, a seconda che occupi posizione di ordine pari o posizione di ordine dispari, viene convertito in un valore numerico, in base alle tabelle di corrispondenza successivamente riportate...

Per i sette caratteri con posizione di ordine

- pari: file CFTabellaPari.tsv
- dispari: file CFTabellaDispari.tsv

Generatore codice fiscale: Carattere alfabetico di controllo

- I valori numerici così determinati vengono addizionati e la somma si divide per il numero 26.
- Il carattere di controllo si ottiene convertendo il resto di tale divisione nel carattere alfabetico ad esso corrispondente nel file CFTabellaResto.tsv

Cavia

```
nome = 'alessandro'
cognome = 'manzoni'
dataNascita = '07031785'
genere = 'M'
comune = 'milano'
MNZLSN85C07F205D
```

- I cognomi che risultano composti da più parti o comunque separati od interrotti, vengono considerati come se fossero scritti secondo un'unica ed ininterrotta successione di caratteri.
- Per i soggetti coniugati di sesso femminile si prende in considerazione soltanto i soggetti coniugati di sesso femminile si prende in considerazione soltanto i
- Se il cognome contiene tre o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il cognome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il cognome contiene una consonante e due vocali, si rilevano, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il cognome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il cognome e' costituito da due sole vocali, esse si rilevano, nell'ordine, e si assume come terzo carattere la lettera x (ics).

- I cognomi che risultano composti da più parti o comunque separati od interrotti, vengono considerati come se fossero scritti secondo un'unica ed ininterrotta successione di caratteri.
- Per i soggetti coniugati di sesso femminile si prende in considerazione soltanto i
- Se il cognome contiene tre o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il cognome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il cognome contiene una consonante e due vocali, si rilevano, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il cognome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il cognome e' costituito da due sole vocali, esse si rilevano, nell'ordine, e si assume come terzo carattere la lettera x (ics).

consonanti + voca<u>li + x</u>

- I cognomi che risultano composti da più parti o comunque separati od interrotti, vengono considerati come se fossero scritti secondo un'unica ed ininterrotta successione di caratteri.
- Per i soggetti coniugati di sesso femminile si prende in considerazione soltanto i compane da publica
- Se il cognome contiene tre o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il cognome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il cognome contiene una consonante e due vocali, si rilevano, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il cognome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il cognome e' costituito da due sole vocali, esse si rilevano, nell'ordine, e si assume come terzo carattere la lettera x (ics).

```
consonanti + vocali + x
stringa[:3]
```

```
def estraiCarCognome(parola):
    parola = parola.replace(' ', '')
    parola = parola.upper()
    frammentoConsonanti = ''
    for carattere in parola:
        if carattere not in ['A', 'E', 'I', 'O', 'U']:
            frammentoVocali += carattere
    parolaRiordinata = frammentoConsonanti + frammentoVocali +
    'XXX'
    return parolaRiordinata[:3]
```

- I nomi doppi, multipli o comunque composti, vengono considerati come scritti per esteso in ogni loro parte e secondo un'unica ed ininterrotta successione di caratteri.
- Se il nome contiene quattro o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la terza e la quarta consonante.
- Se il nome contiene tre consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il nome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il nome contiene una consonante e due vocali, i tre caratteri da rilevare sono, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il nome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il nome e' costituito da due sole vocali, esse si rilevano nell'ordine, e si assume come terzo carattere la lettera x (ics).

- I nomi doppi, multipli o comunque composti, vengono considerati come scritti per esteso in ogni loro parte e secondo un'unica ed ininterrotta successione di caratteri.
- Se il nome contiene quattro o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la terza e la quarta consonante.
- Se il nome contiene tre consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il nome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il nome contiene una consonante e due vocali, i tre caratteri da rilevare sono, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il nome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il nome e' costituito da due sole vocali, esse si rilevano nell'ordine, e si assume come terzo carattere la lettera x (ics).

consonanti + vocali + x

- I nomi doppi, multipli o comunque composti, vengono considerati come scritti per esteso in ogni loro parte e secondo un'unica ed ininterrotta successione di caratteri.
- Se il nome contiene quattro o più consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la terza e la quarta consonante.
- Se il nome contiene tre consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima, la seconda e la terza consonante.
- Se il nome contiene due consonanti, i tre caratteri da rilevare sono, nell'ordine, la prima e la seconda consonante e la prima vocale.
- Se il nome contiene una consonante e due vocali, i tre caratteri da rilevare sono, nell'ordine, quella consonante e quindi la prima e la seconda vocale.
- Se il nome contiene una consonante e una vocale, si rilevano la consonante e la vocale, nell'ordine, e si assume come terzo carattere la lettera x (ics).
- Se il nome e' costituito da due sole vocali, esse si rilevano nell'ordine, e si assume come terzo carattere la lettera x (ics).

```
consonanti + vocali + x
stringa[1,3,4] or stringa[:3]
```

```
def estraiCarNome(parola):
    parola = parola.replace(' ', '')
    parola = parola.upper()
    for carattere in parola:
        if carattere not in ['A', 'E', 'I', 'O', 'U']:
            frammentoConsonanti += carattere
            frammentoVocali += carattere
    if len(frammentoConsonanti) > 3:
            0] + frammentoConsonanti[2] + frammentoConsonanti[3]
    'XXX'
    # print parolaRiordinata
    return parolaRiordinata[:3]
```

Data

- I due caratteri numerici indicativi dell'anno di nascita sono, nell'ordine, la cifra delle decine e la cifra delle unità dell'anno stesso.
- Il carattere alfabetico corrispondente al mese di nascita e' quello stabilito per ciascun mese nella seguente tabella:

Gennaio = A	Maggio = E	Settembre = P
Febbraio = B	Giugno = H	Ottobre = R
Marzo = C	Luglio = L	Novembre = S
Aprile = D	Agosto = M	Dicembre = T

Data

Sesso

- I due caratteri numerici indicativi del giorno di nascita e del sesso vengono determinati nel modo seguente:
 - Per i soggetti maschili il giorno di nascita figura invariato, con i numeri da uno a trentuno, facendo precedere dalla cifra zero i giorni del mese dall'uno al nove.
 - Per i soggetti femminili il giorno di nascita viene aumentato di quaranta unità, per cui esso figura con i numeri da quarantuno a settantuno.

```
def estraiCarGiorno(parola, carGenere):
    if carGenere == 'F':
        giorno = int(parola) + 40
        print giorno
        parola = str(giorno)
    return parola
```

Luogo

 I quattro caratteri alfanumerici indicativi del comune italiano o dello Stato estero di nascita, sono costituiti da un carattere alfabetico seguito da tre caratteri numerici, secondo la codifica stabilita dall'Agenzia del Territorio e contenuta nel file CodiceComuniItalia.csv

```
CODICI COMUNI D'ITALIA - COMUNI ATTUALI alla data del 6 febbrai Codice; Prov; Denominaz Italiana; Denominaz Estera; Cod Cat; Uff cat A154; VI; ALBETTONE;; D7AC; VI; VI; VI00; Vicenza; 24002;; A155; CZ; ALBI;; T2AC; CZ; CZ00; Catanzaro; 79002;; A157; TO; ALBIANO D'IVREA;; A1AF; TO; TO; TO40; Ivrea; 1004;; A158; TN; ALBIANO;; E1AD;;;;; 22002;; A159; MB; ALBIATE;; C1AI; MI; MI; MI20; Milano 2; 108003;; A160; CS; ALBIDONA;; T3AH; CS; CS; CS00; Cosenza; 78006;; A161; PD; ALBIGNASEGO;; D3AD; PD; PD; PD00; Padova; 28003;;
```

Luogo

```
def estraiCarComune(parola):
    dizComuni = {}
    nomeFile = 'CFCodiceComuniItalia.csv'
    fileComuni = open(nomeFile, 'r')
    listaComuni = fileComuni.readlines()
    fileComuni.close()
    for linea in listaComuni:
        frammenti = linea.split(';')
        # print frammenti
        dizComuni[frammenti[2]] = frammenti[0]
    return dizComuni[parola.upper()]
```

Carattere alfabetico di controllo

Il sedicesimo carattere ha funzione di controllo dell'esatta trascrizione dei primi quindici caratteri e viene determinato in questo modo:

 ciascuno degli anzidetti quindici caratteri, a seconda che occupi posizione di ordine pari o posizione di ordine dispari, viene convertito in un valore numerico, in base alle tabelle di corrispondenza successivamente riportate...

Per i sette caratteri con posizione di ordine

```
    pari: file CFTabellaPari.tsv
```

• dispari: file CFTabellaDispari.tsv

CFTabellaPari.tsv

#Cai	ratte	re	Valo	re	Carat	ttere	Valore	c Carattere
	Valo	re	Cara	ttere	V	alore		
0	0	9	9		8		17	
1	1	Α	0	J	9		18	
2	2		1		10		19	
3	3	С	2		11	U	20	
4	4	D	3		12		21	
5	5		4		13	W	22	
6	6		5	0	14		23	
7	7		6		15		24	
8	8	Н	7	Q	16		25	

CFTabellaDispari.tsv

#Ca	ratter		Valor		Carat	tere	Valor	e Carattere
	Valor		Carat	tere	Va	lore		
0	1	9	21		19		8	
1	0	Α	1	J	21		12	
2	5		0		2		14	
3	7	С	5		4	U	16	
4	9	D	7		18		10	
5	13		9		20	W	22	
6	15		13	0	11		25	
7	17		15		3		24	
8	19	Н	17	0	6	Z	23	

Controllo

```
def estraiCarControllo(parola):
    file = open('CFTabellaDispari.tsv', 'r')
    listaDispari = file.readlines()
    file.close()
    file = open('CFTabellaPari.tsv', 'r')
    listaPari = file.readlines()
    file.close()
    del listaDispari[0]
    del listaPari[0]
    dizDispari = {}
    for linea in listaDispari:
        frammenti = linea.strip().split('\t')
        for i in range(4):
            dizDispari[frammenti[2 * i]] = frammenti[2 * i + 1]
    dizPari = {}
    for linea in listaPari:
        frammenti = linea.strip().split('\t')
        for i in range(4):
            dizPari[frammenti[2 * i]] = frammenti[2 * i + 1]
```

Carattere alfabetico di controllo

- I valori numerici così determinati vengono addizionati e la somma si divide per il numero 26.
- Il carattere di controllo si ottiene convertendo il resto di tale divisione nel carattere alfabetico ad esso corrispondente nel file CFTabellaResto.tsv

CFTabellaResto.tsv

```
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
```

Controllo

```
file = open('CFTabellaResto.tsv', 'r')
listaResto = file.readlines()
file.close()
dizResto = {}
     frammenti = linea.strip().split('\t')
     dizResto[frammenti[0]] = frammenti[1]
# print dizResto
i = 0
somma = 0
for el in parola:
     if i % 2 == 1: # attenzione i è indice: 0 = primo
         somma += int(dizPari[el])
         print dizPari[el]
         somma += int(dizDispari[el])
         print dizDispari[el]
     i += 1
print 'somma = ', somma, 'resto = ', somma % 26
car = dizResto[str(somma % 26)]
```

```
nome = raw input('inserire il nome ')
cognome = raw_input('inserire il cognome ')
dataNascita = raw_input('inserire la data di nascita nel formato
    ggmmaaaa ')
genere = raw_input('inserire il genere M o F ')
comune = raw_input('inserire il comune di nascita ')
nome = nome.replace(' ', '')
cognome = cognome.replace(' ', '')
giorno = dataNascita[:2]
mese = dataNascita[2:4]
anno = dataNascita[6:8]
print anno, "\n"
carCognome = estraiCarCognome(cognome)
carNome = estraiCarNome(nome)
carMese = estraiCarMese(mese)
carGiorno = estraiCarGiorno(giorno, genere)
carComune = estraiCarComune(comune)
CF = carCognome + carNome + anno + carMese + carGiorno +
carControllo = estraiCarControllo(CF)
```

Dall'RNA alle proteine

I moderni strumenti di laboratorio permettono di estrarre le sequenze di nucleotidi (RNA) espresse in un certo individuo. Queste informazioni vengono salvate in file di formato standard FASTA. Sapendo che i nucleotidi vengono letti in triplette (codoni) e che ciascuna di queste triplette corrisponde ad un ammino-acido nella sequenza proteica, si chiede di scrivere un programma che abbia come input il file fasta A06662-RNA. fasta e la tabella di conversione codonTable e restituisca le possibili sequenze proteiche. Inoltre, per chiarezza, si sostituiscano gli ammino-acidi 'STOP' con il carattere '*'

A06662-RNA.fasta

>A06662.1 Synthetic nucleotide sequence of the human GSH transferase pi gene

codonTable

```
GCA
. . .
```

fastaRNA2Prot.py

```
with open('codonTable', 'r') as codons:
    rna2cod = codons.readlines()
    codons.close()
# print rna2cod
dizRNA2Prot = {}
    tokens = line.strip().split('\t')
    dizRNA2Prot[tokens[0]] = tokens[1]
with open('A06662-RNA.fasta', 'r') as rna:
    rnaFile = rna.readlines()
    rna.close()
del rnaFile[0]
# print rnaFile
rnaString = ''
    rnaString += line.strip()
```

fastaRNA2Prot.py

```
prots = []
for readFrame in range(3):
    for i in range(readFrame, len(rnaString) - readFrame - 1,
    3):
        codon = rnaString[i:i + 3]
        protFrag = dizRNA2Prot[codon]
        if protFrag == 'STOP':
            prot += '*'
            prot += protFrag
    prots.append(prot)
for prot in prots:
```

uenaacfarsaos???





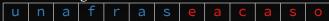


unafras eacaso





si decide il taglio a metà frase





si decide il taglio a metà frase



si decide allineano i 2 frammenti

u	n	а	f	r	а	S
е	а	С	а	S	0	



si decide il taglio a metà frase



si decide allineano i 2 frammenti



si scrive una lettera da un frammento di seguita ad una dell'altro





si decide il taglio a metà frase



si decide allineano i 2 frammenti



si scrive una lettera da un frammento di seguita ad una dell'altro





si decide il taglio a metà frase



si decide allineano i 2 frammenti



si scrive una lettera da un frammento di seguita ad una dell'altro

u e n a a c f a r s a o s



si decide il taglio a metà frase



si decide allineano i 2 frammenti



si scrive una lettera da un frammento di seguita ad una dell'altro

u e n a a c f a r s a o s

```
messaggio = raw_input('inserisci il messaggio ')
# CODTFTCA
messaggioT = messaggio.replace(' ', '')
iSpezza = (len(messaggioT) + 1) / 2
print 'len = ', len(messaggioT), 'iSpezza', iSpezza
stringa1 = messaggioT[:iSpezza]
stringa2 = messaggioT[iSpezza:]
print stringa1, len(stringa1)
print stringa2, len(stringa2)
lMessaggioCifrato = list(messaggioT)
for i in xrange(iSpezza):
    indice = i * 2
    lMessaggioCifrato[indice] = stringa1[i]
    print '', indice + 1
    if i < len(stringa2):</pre>
        lMessaggioCifrato[indice + 1] = stringa2[i]
messaggioCifrato = ''.join(lMessaggioCifrato)
```

```
for i in xrange(len(messaggioCifrato)):
    if i % 2 == 0:
        frammento1 += messaggioCifrato[i]
        frammento2 += messaggioCifrato[i]
print frammento1, frammento2
```