

Ejercicio 1 Python 3.x

Objetivo: Repaso de estructuras básicas de python (listas, tuplas, diccionarios, ...) y su sintaxis.

Referencias: <http://www.python.org/>

Dada la cadena:

cadena = "El/DT perro/N come/V carne/N de/P la/DT carnicería/N y/C de/P la/DT nevera/N y/C canta/V el/DT la/N la/N la/N ./Fp"

- 1) Obtener un diccionario, que para cada categoría, muestre su frecuencia. Ordenar el resultado alfabéticamente por categoría.

```
C 2
DT 4
Fp 1
N 7
P 2
V 2
```

- 2) Obtener un diccionario, que para cada palabra, muestre su frecuencia, y una lista de sus categorías morfosintácticas con su respectiva frecuencia. Mostrar el resultado ordenado alfabéticamente por palabra.

```
. 1 Fp 1
canta 1 V 1
carne 1 N 1
carnicería 1 N 1
come 1 V 1
de 2 P 2
el 2 DT 2
la 5 DT 2 N 3
nevera 1 N 1
perro 1 N 1
y 2 C 2
```

- 3) Calcular la frecuencia de los todos los bigramas de la cadena, considerar un símbolo inicial <S> y un símbolo final </S> para la cadena.

```
('DT', 'N') 4
('N', 'V') 1
('N', 'C') 2
('N', 'Fp') 1
('N', 'N') 2
('C', 'V') 1
('V', 'N') 1
('V', 'DT') 1
('P', 'DT') 2
('Fp', '</S>') 1
('<S>', 'DT') 1
('C', 'P') 1
('N', 'P') 1
```

- 4) Construir una función que devuelva las probabilidades léxicas $P(C|w)$ y de emisión $P(w|C)$ para una palabra dada (w) para todas sus categorías (C) que aparecen en el diccionario construido anteriormente. Si la palabra no existe en el diccionario debe decir que la palabra es desconocida.

Por ejemplo, para la palabra $w="la"$, debería devolver:

$P(DT | la) = 0.400000$

$P(N | la) = 0.600000$

$P(la | DT) = 0.500000$

$P(la | N) = 0.428571$