

## 8. házi feladat

### Megoldás

Ferenczi Zsanett

#### 1. A feladat

Írj egy konvertert az IOB és a BIE1 taggelés között!

#### 2. A megoldás

##### 2.1. IOB taggelés

- I a szekvencia belsejét,
- B a szekvencia kezdő elemét,
- O a szekvencián kívüli elemeket jelöli.

##### 2.2. BIE1 taggelés

- B a szekvencia kezdő elemét,
- I a szekvencia belsejét,
- E a szekvencia utolsó elemét,
- 1 az egyelemű szekvenciákat,
- O a szekvencián kívüli elemeket jelöli.

## 2.3. Megoldás Python-ban

Kétféleképp is meg lehet közelíteni a feladatot: a függvénynek átadott tagek listáján végighaladva egy új listához fűzzük az átalakított címkéket, vagy a paraméterként kapott listát lemásoljuk, és csak azokat az elemeket cseréljük le, amik nem egyeznek meg a két taggelés között.

### 2.3.1. O és B kezdetű címkék

Mindkét esetben fontos megjegyezni, hogy az O címkék, és a B-vel kezdődő címkék azonosak a két taggelés között, azokat tehát nem alakítjuk át, míg az I-vel kezdődő címkénél több dologra is figyelniünk kell, amikor  $\text{IOB} \rightarrow \text{BIE1}$  konvertert írunk.

### 2.3.2. I kezdetű címkék átalakítása

- amikor az I kezdetű címke két I között vagy egy B után és I előtt áll (ugyanazzal a végződéssel, pl. `ORG`), akkor biztosak lehetünk benne, hogy belső elemről van szó, tehát marad az adott tag

pl. `B-ORG I-ORG I-ORG`  $\rightarrow$  `B-ORG I-ORG E-ORG`

- amikor egy I tag előtt B vagy azonos I elem áll, de utána nem ugyanaz következik, akkor a szekvencia végére értünk, tehát le kell cserélni az elemet E-re

pl. `B-ORG I-ORG I-PERS`  $\rightarrow$  `B-ORG E-ORG 1-PERS`

- egyéb I tag esetén egyelemű szekvenciáról lehet csak szó, így ezt 1-gyel kezdődő elemre cseréljük

pl. `B-PERS I-PERS I-ORG O`  $\rightarrow$  `B-PERS E-PERS 1-ORG O`

A Python kód megtalálható a [githubon](#).