

Annotációs útmutató az emMorph címkék javításához

A feladat: a zéró létigék és az elliptált igék a függőségi elemzéssel ellátott mondatokba.

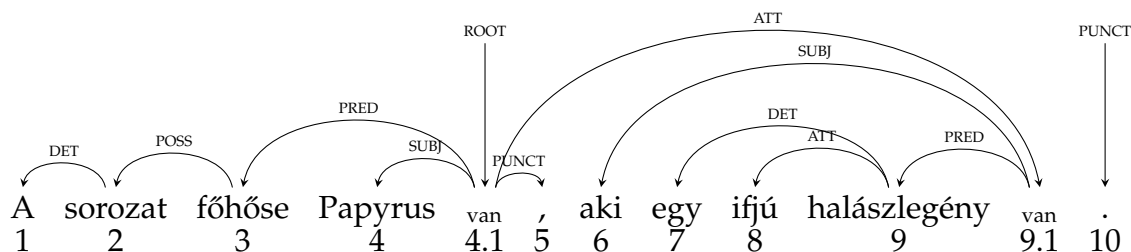
A feladatot az Annotatrix¹ felületén kell elvégezni. A mondatot, amelybe a zéró igét kell beilleszteni, másoljuk be a felületen lévő középső nagy szövegdobozba. Jobb oldalt felül a CoNLL-U formátum legyen kiválasztva. Ezután megjelenik a mondat elemzési fája.

A mondatba úgy kell beszúrni egy új szót, hogy a CoNLL-U forrásba a megfelelő helyre beillesztünk egy plusz sort. A zéró igéhez tartozó tövet, szófajt és morfológiai jegyeket kézzel kell megadni. Az indexek esetleges újrakiosztását az Annotatrix automatikusan elvégzi.

A zéró igék vonzatai a zéró igéhez kerülnek. Ehhez vagy kézzel kell behúzni a függőségi fában a megfelelő éleket, vagy be kell írni a CoNLL-U forrásba.

1. Zérólétigék

A zérólétigéket oda illesszük be, ahol az a mondat múlt idejű vagy feltételes módú változatában megjelenne. A zéró létigék tehát új tokenként kerülnek a fájlba arra a helyre, ahol múlt időben testes létigeként jelennének meg, saját kombinált indexet kapnak, ami a zéró létigét megelőző elem ID-jéből képződik. Az 1. példában egy olyan mondat látható, ahol a függőségi fába két zéró létigét is be kellett illeszteni.



1. ábra. Az összetett mondat fölérendelt tagmondatának zéró létigéje alá van rendelve az alárendelt mellékmondat zéró létigéje. A második sorban a kiosztott indexek láthatók a zéró létigék kombinált indexével együtt.

A 2. ábrán a fenti példa CoNLL-U forrása látható.

2. Elliptált igék

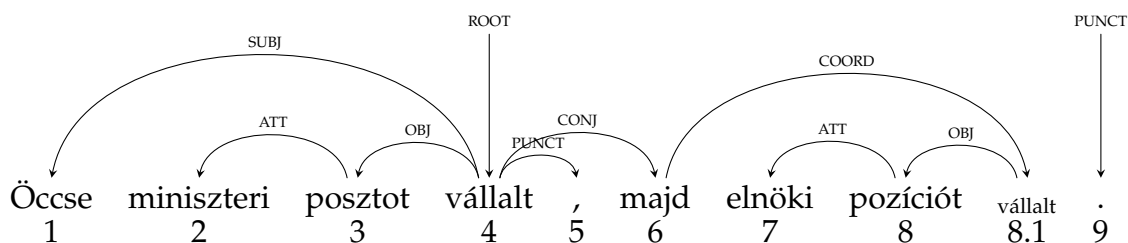
Az elliptált igéket oda illesszük be, ahol az megjelenne, ha kiejtenénk.

¹<https://github.com/jonorthwash/ud-annotatrix>

1	A	a	DET	DET	Definite=Def PronType=Art	2	DET	-	-										
2	sorozat	sorozat	NOUN	NOUN	Case=Nom Number=Sing	3	POSS	-	-										
3	főhőse	főhős	NOUN	NOUN	Case=Nom Number=Sing Number[psor]=Sing Person[psor]=3	4.1	PRED	-	-										
4	Papyrus	Papyrus	PROPN	PROPN	Case=Nom Number=Sing	4.1	SUBJ	-	-										
4.1	KOPULA	van	VERB	VERB	Definite=Ind Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin Voice=Act	0	root	-	-										
5	,	,	PUNCT	PUNCT	-	4.1	PUNCT	-	-										
6	aki	aki	PRON	PRON	Case=Nom Number=Sing Person=3 PronType=Rel	9.1	SUBJ	-	-										
7	egy	egy	DET	DET	Definite=Ind PronType=Art	9	DET	-	-										
8	ifjú	ifjú	ADJ	ADJ	Case=Nom Degree=Pos Number=Sing	9	ATT	-	-										
9	halászegény	halászegény	NOUN	NOUN	Case=Nom Number=Sing	9.1	PRED	-	-										
9.1	KOPULA	van	VERB	VERB	Definite=Ind Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin Voice=Act	4.1	ATT	-	-										
10	.	.	PUNCT	PUNCT	-	0	root	-	-										

2. ábra. Egy példamondat két beillesztett zérólétigével.

Az elliptált ige a zero létigéhez hasonló, kombinált indexet kap. A 3. példában egy olyan mondat látható, ahol a függőségi fába egy elliptált igét kellett beilleszteni.



3. ábra. A mellérendelés első mondatában szereplő ige a második mondatban testetlenül van jelen. Ezért egy zero alakot illesztettük be a függőségi fába, így a második mondatban szereplő vonzat már kapcsolódni tud a saját testetlen igéjéhez.

A 4. ábrán a fenti példa CoNLL-U forrása látható.

1	öccse	öcs	NOUN	NOUN	Case=Nom Number=Sing Number[psor]=Sing Person[psor]=3	4	SUBJ	-	-										
2	miniszteri	miniszteri	ADJ	ADJ	Case=Nom Degree=Pos Number=Sing	3	ATT	-	-										
3	posztot	poszt	NOUN	NOUN	Case=Acc Number=Sing	4	OBJ	-	-										
4	vállalt	vállal	VERB	VERB	Definite=Ind Mood=Ind Number=Sing Person=3 Tense=Past VerbForm=Fin Voice=Act	0	root	-	-										
5	,	,	PUNCT	PUNCT	-	4	PUNCT	-	-										
6	majd	majd	CONJ	CONJ	-	4	CONJ	-	-										
7	elnöki	elnöki	ADJ	ADJ	Case=Nom Degree=Pos Number=Sing	8	ATT	-	-										
8	pozíciót	pozíció	NOUN	NOUN	Case=Acc Number=Sing	8.1	OBJ	-	-										
8.1	ZÉRÓ	vállalt	vállal	VERB	VERB	Definite=Ind Mood=Ind Number=Sing Person=3 Tense=Past VerbForm=Fin Voice=Act	6	COORD	-	-									
9	.	.	PUNCT	PUNCT	-	0	root	-	-										

4. ábra. Egy példamondat egy elliptált igével.