Predstavitev označevanja koreferenčnosti

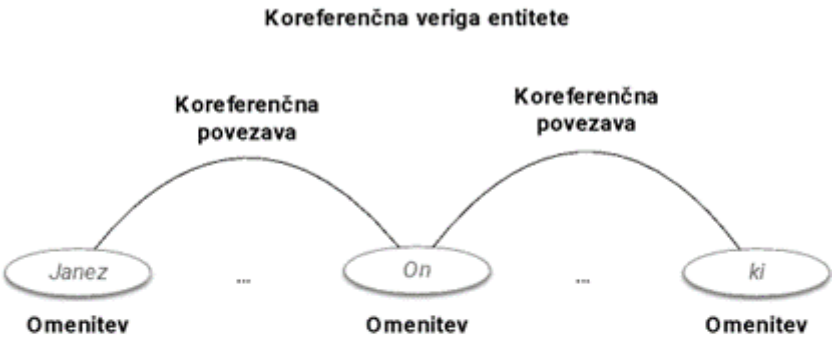
V tem poglavju je strnjeno predstavljeno označevanje koreferenc. Podrobnejšo predstavitev najdete v smernicah v poglavju Označevalne smernice.

Dele besedila, ki se nanašajo na istega referenta, imenujemo **omenitve** (angl. *mentions*). Omenitve se lahko pojavljajo preko različnih stavkov, povedi, odstavkov ali drugih delov besedila. Pri označevanju jih povežemo v t. i. **koreferenčno verigo** (angl. *coreference chain*). Koreferenčne verige nakazujemo z različnimi barvami, kot na primer v naslednjih povedih:

- **[1.a]** Peter ima dva psa.
- **[1.b]** On se velikokrat igra z njima.

Omenitvi *Peter* in *On* se nanašata na istega referenta, tj. isto konkretno osebo, kar pomeni, da sta koreferenčni. Enako velja za omenitvi *njima* in *dva psa*, ki se nanašata na isti par živali.
Naloga pri označevanju koreferenčnosti je identifikacija in povezovanje omenitev. Označuje se samo omenitve, ki so koreferenčne z drugo omenitvijo. Delov besedila, ki niso koreferenčni z nobenim drugim delom besedila, torej ne označujemo.

Označevanja koreferenčnosti prikazuje še spodnja shema. Prikazane so tri omenitve, ki se sklicujejo na določeno entiteto. Povezave med zaporednimi omenitvami, ki se sklicujejo na isto entiteto, so koreferenčne povezave in skupaj z vsemi omenitvami tvorijo koreferenčno verigo. Vsaka omenitev ima določene še lastne oznake.



Dodante oznake omenitev:

Type
Copular-N, Copular-S, Appositive, Zero anaphora, Clause, Regular
Referent *(* only for Copular-S, otherwise NA)*
Subject/Object/NA
Discontinuous
True/False
Ellipsis
True/False

Revision #4
Created 2023-03-03 08:12:19 UTC by Tina Munda
Updated 2023-11-14 19:14:32 UTC by Tina Munda