

Report tree_projection

In this assignment, I implemented the tree projection and tested it on 3 language pairs:

- En-ru
- Ru-cs
- Fr-en

The projection is done in the following way:

- 1) I got the source sentence and retrieved all subtrees from there (in the format 'head: [dependencies]').
- 2) I looped through all heads in source and projected them to the target using union alignment. If source head had more than one aligned word in target, I chose the one with the same UPOS, or just the first one, if no UPOS are the same. All other aligned words became dependents for a chosen head.
- 3) For each source head I looped through all its dependencies and projected them as well, using the alignment.
- 4) In steps 2 and 3 I also projected the deprels either using alignment, or using the separate function I've created, which assigns the target deprel based on UPOS of the target word.
- 5) After that, I looped through target words to fill the gaps in deprels with the same function again.
- 6) Following the projection, I've implemented the rooted trees check:
 - a) I got all roots in the sentence: if there is more than one, I chose the one with the highest number of dependents and turned other roots into dependents of it; if there is no root, I chose the head with the highest number of dependents and turned it into the root.
 - b) For all nodes without head, I put the root as a head.
 - c) I checked cycles – from each node I constructed the path from it to the root and if there is a cycle, I destroyed it by redirecting the cycling node to the root.

The resulting accuracy is the following:

Language pair	Head	Deprel	LAS
En-ru	0.33	0.45	0.20
Ru-cs	0.34	0.48	0.23
Fr-en	0.40	0.56	0.30

It is not much higher than the accuracy in the basic approach, where I tried to simply copy the deprels and heads from the source using alignment:

Language pair	Head	Deprel	LAS
En-ru	0.28	0.40	0.19
Ru-cs	0.28	0.43	0.20
Fr-en	0.37	0.53	0.28