Quality measures

Exercise 3.1. Consider the following two dependency structures for the sentence “Für 200 Mark können Kinder zwischen neun und zwölf Jahren mitfahren.” from the NEGRA corpus. The first one is obtained by conversion from the gold standard. The latter was produced by a dependency parser.

Calculate the UAS and the LAS. In practice, you have to compute one score for the entire test corpus. Develop different ways to do this and discuss potential advantages/disadvantages.

Exercise 3.2. Consider the gold tree (top) from the TIGER corpus and the system output (bottom).
Volks- und Raiffeisenbanken zeichnen innovative Beschäftigungskonzepte im Mittelstand aus.

Compute labeled precision, labeled recall, labeled accuracy and the (labeled) F-measure for this pair of trees. In advance, specify the set of instances, the set of relevant items, and the set of selected items.

Additional exercise 3.3. Use disco-dop¹ to compute labeled precision, recall, and F1 for the above tree pair.

¹https://github.com/andreasvc/disco-dop