A POS Tagger Comparison -
Three POS Taggers compared individually and combined on Prose Text and Poetry

Matteo Badaracco

Master thesis in Computer Science

With this work we tried to create two a state-of-the art POS taggers and compare them on prose and poetry. Therefore, our contributions comprise a CRF model and an LSTM model trained on prose. A comparison is done between these two POS taggers and a TreeTagger already trained on prose and poetry. We also tried to combine them in order to see if it was possible to get better performance in combining multiple POS taggers. Finally, we created an original, manually annotated poetry corpus ready for testing or training some POS taggers.

Prof. Rolf Ingold