

## 4.5.21 exercises NLTK chap 2

ex 1. nltk text generation

Train an ngram model on the english bible and get it to generate random text. Experiment with different start words and window lengths.

How intelligible is the text?

Discuss the strengths and weaknesses of this method of generating random text and propose alternatives.

## 4.5.21 exercises NLTK chap 2

ex 2. wn meronyms

Use the BROWN religion vs. romance subcorpora.

Which one has more NOUN meronyms? Remember that there are two types of meronyms.

## 4.5.21 exercises NLTK chap 2

ex 3. wn polysemy

- a) The polysemy of a word is the number of senses it has. Using the BROWN religion vs. romance subcorpora, which subcorpus is generally more polysemous or 'ambiguous'?
- b) Compute the average polysemy for the religion NOUNS in a 2 liner!

## 4.5.21 exercises NLTK chap 2

ex 4. wn word similarity

wn provides 6 different measures of word similarity.

Which of these measures predicts the human ratings from the 'simlex999' gold standard best?