

# Tools and Techniques for Speech and Language Processing

University of Illinois at Urbana-Champaign

Week 09 of 16 — Day 17 of 29

# Quiz (~10 minutes)

# Go over hw09

# Questions on hw10

# Finding synonyms with nltk

- Get the raw text of 'English-Latin1' from `nltk.corpus.udhr`.
- Split the raw text into sentences using the `split()` function.
- Tokenize each sentence using `nltk.word_tokenize()`.
- For each token, find a synonym - i.e. the name of the first lemma of the first synset in wordnet of which the token is a member.
- Print an output string for each sentence, where each token has been replaced by a synonym.
- If no synonym can be found, the original token should be printed.

You will find it helpful to look through [http://www.nltk.org/\\_modules/nltk/corpus/reader/wordnet.html](http://www.nltk.org/_modules/nltk/corpus/reader/wordnet.html)