

# Tools and Techniques for Speech and Language Processing

University of Illinois at Urbana-Champaign

Week 08 of 16 — Day 15 of 29

# Quiz (~10 minutes)

# Go over hw08

# Questions on hw09

Retrieve the html from one of the following webpages:

- `http://computational.linguistics.illinois.edu/ling402/Fall2017/index.html`
- `http://computational.linguistics.illinois.edu/ling402/Fall2017/schedule.html`

Use Python regular expressions (re module) to retrieve certain elements (like tags or text) from the html. For example...

- Retrieve the names of all tags in the html.
- Retrieve the names of all tags with corresponding end tags.
- Retrieve all tags with attributes and create a dictionary that stores a dictionary of attributes at each tag name.