Introductory lecture

Introductions

- Instructor
- Practicum leader
- Students

Syllabus review

(Actually read the policies out loud.)

Attendancy policy

- In-person class will happen as indicated on schedule, barring illness, injury, or inclement weather.
- Please attend class.

Resources

- Listed office hours (but please write ahead of time to schedule a slot)
- The computational linguistics lab (rm. 7400.13)
- The `cuny-compling` Slack; ask Kyle for an invite

Why this class

This class sits at the intersection of three topics:

1. Best practices in software development: Git and GitHub, command-line interfaces, unit tests, "tooling"
2. Mathematical linguistics and grammar engineering: finite automata, rational relations, context-free grammars, and the Pynini finite-state toolkit
3. Machine learning: naïve Bayes, logistic regression, regularization, TF-IDF, hyperparameter tuning, and the Scikit-learn classification toolkit

These topics have strong synergies, and even if they didn't, you need to know all three to be a working computational linguist. This is in contradistinction to Methods I, where most of the material is just relevant to working in Python in the sciences in general.