

Final Review

L435/L555

Fall 2016

1 Organization

The exam will be take-home, starting Thursday, December 8 after class and due Thursday, December 15 at 5pm, via email.

You are to work on your own, unassisted, though you can (and should) ask me for clarifications, where needed.

2 Format

You will be asked to write different short programs. You may also be asked to correct a program which is not written properly.

3 Topics

Go over the slides (& the book & any other resources) and make sure you understand the topics listed below. The final is not strictly cumulative, in that it focuses only on topics since the midterm, but as you know, programming is a cumulative type of process!

- NLTK
 - Text data types & Text methods (`.concordance()`, `.similar()`, `.common_context()`, `.collocations()`)
 - Use of `set()` (cf. word type counts)
 - `FreqDist`, `ConditionalFreqDist`
 - `bigrams()` function
 - `nltk.Text()` conversion
 - `word_tokenize()`
 - Stemmers & lemmatizers
 - Accessing (tagged) corpora (e.g., Brown corpus)
 - `sent_tokenize()`

- Dictionaries
 - Core idea: key-value mapping
 - Basic building & accessing (e.g., `in`, `get()`)
 - Methods (`.values()`, `.keys()`, `.items()`, `.copy()`, `.update()`, `.pop()`)
 - `del`
 - Sorting by key & by value
 - Embedding lists & dictionaries within dictionaries
- Functions
 - Making code more reusable
 - Parameters/Arguments (including defaults)
 - Positional, keyword, & collected parameters
 - `return`
 - Scope of variables
 - Recursion (and functions calling functions more generally)
- Modules
 - `import X`, `from X import *`
 - Main code vs. module
 - Finding modules & information on modules
- Classes
 - Objects & classes
 - Methods & attributes
 - Superclasses
- Regular expressions
 - Basic RE syntax (e.g., using `egrep` outside Python)
 - `.compile()`, `.search()`, `.match()`, `.split()`, `.findall()`, `.sub()`
 - Pattern objects & `.group()`