

## Checklist

### General

- ☐ Checked your repository at `github.uiowa.edu` to make sure it has the version of the files you expected
- ☐ Filled in `PROGRESS_REPORT_APRIL24.txt` on time
- ☐ Filled in `PROGRESS_REPORT_MAY1.txt` on time

### Code

- ☐ Command `help` prints the help message for all of your commands
- ☐ Command `index FILE` updates the sentences and vectors to include the contents of `FILE`, where `FILE` is the filename for the input text
- ☐ Command `sentences` prints the sentences and number of sentences
- ☐ Command `vectors` prints the vectors
- ☐ Command `topj Q J` will run the Top-J query: find the `J` words most similar to word `Q`
- ☐ Improved the program's running time
- ☐ Command `measure SIM`, where `SIM` can be `cosine`, `euc`, or `eucnorm`. Will affect which similarity measure is used for subsequent `topj` commands.
- ☐ JUnit tests for `euc` and `eucnorm`.
- ☐ Command `kmeans k iters` runs K-means clustering with `k` clusters and `iters` number of iterations.

### Writing in project.pdf

- ☐ Answered a,b,c,d (Part 4)
- ☐ Report findings on Top-J comparing indexing one text to indexing two texts (Part 5)
- ☐ Report findings on Top-J with three similarity measures (Part 5)
- ☐ Plot of average distance vs. iteration; how many iterations to converge? (Part 6)