

Fall Semester 2017
Iowa State University

LING 120 - Language and Computers

Midterm Project Descriptions
Submission Deadline: 11 October 2017

General Instructions: You have to work in groups of 4 (assigned by me), pick any **one question** among all the listed ones (e.g., 1 (a), 1(b), 1(c) are three questions, not one!), and present on that. Presentations have to be 10 minutes long, with 5 minutes for questions. Carries 20 marks. It is okay if 2 teams pick the same topic.

1. Questions related to Topic 2:
 - (a) Describe how to calculate edit distance between two words (text-book has enough information, and I gave some video link in the class), and how it is useful for doing spelling correction.
 - (b) Describe how the ideas such as probability of language, and n-grams are useful for doing spelling and grammar correction (text-book has it, again).
 - (c) Use the web demo of LanguageTool.org for any two variations of English (e.g., American-British) it supports, and present an evaluation of the tool in terms of its correctness, how it gives feedback etc. Discuss also about how the tool works.
2. Questions related to Topic 3:
 - (a) Use DuoLingo, take a couple of lessons for any **two** languages, preferably with different writing systems. Present about the sequence of exercises, about differences between the two languages in terms of practice etc. Also add comments on what you think of using DuoLingo instead of a real classroom setup. Discuss how Duolingo may be evaluating your answers.
 - (b) You have seen the VIEW tool demo in Week 4. Install that tool, work with it for sometime, and present about the working of that, what kind of language processing is for what steps, is it really useful, etc.
3. Questions related to Topic 4:
 - (a) Demonstrate the usefulness of regular expressions for searching text on MS Word. Can we use regular expressions to do search on files in your computer? How? Present about your answers to these questions.

4. Questions covering multiple topics:

- (a) Pick any two of these: chatwithigod.com, mitsuku.com, <http://www.manifestation.com/neurotoys/eliza.php3> and hold a conversation with those programs. Present what you observed, where they work and where they fail, and what is the difference between the two programs you picked. You may want to read a little bit about them on their websites or on Wikipedia.
- (b) Go to: books.google.com/ngrams - what is this website doing? Where is it useful? How does it work? What is required to build something like that? Demonstrate its usefulness through some examples.
- (c) Here is the demo of an online part of speech tagger - http://cogcomp.org/page/demo_view/POS - Present an overview of how it works, how good it is, where it makes errors (e.g., if the text has spelling or grammar errors, will tagging be accurate?) etc. with examples.
- (d) What is this demo doing? - http://cogcomp.org/page/demo_view/Coref. Explain what you understand about the demo (you may have to read some additional documentation)
- (e) Install voice recognition in google chrome browser (<https://goo.gl/swe1NW>) and present a study of how good or bad it is with different voices, accents and different vocabulary. Try to discuss why it is working the way it does, how it can be improved, etc.