Choose ONE essay title and ONE programming problem. The essay length should be ~1000-1500 words. I expect the programming problems to be of a similar level of difficult/effort.

I'm very happy to give tutorial advice on Tuesday afternoons (and read/discuss drafts of essays).

**Essay Titles**

There's an extensive reading list for all the essay titles available from the NLPA webpage. Some of the papers however are only available in books so you will have to visit the main library!

**Syntax and Parsing**

Both Pereira (1985) and Marcus (1980) have proposed parsers that incorporate psycho linguistic observations about how humans process natural language. Describe these parsers, explain what data is incorporated and how it is done. Discuss what you see as the main similarities and differences between the two parsers.

When considering the differences, you should consider at least whether the two approaches make the same predictions concerning human parsing behavior; whether the syntactic rules are separate from the actions of the parser; how much context is used by each parser.

Full references for both papers are available off the webpage.

**Word Senses and Meaning**

"Do I (don't/do) believe in word-senses"

Read Kilgariff's paper "I don't believe in word senses." Do you agree with Kilgariff's argument? Compare it to other authors such as James Pustojovsky. Are word senses discrete psychological categories and what implications does this have for applied Natural Language.

**Discourse Processing**

Text Summarisation is an emerging technology with interesting potential applications. In particular there is an increasing interest in using summarization technologies in Internet Applications. Write an essay which argues the case for a useful application of text summarisation in a realistic real world domain. As part of the essay provide a thorough literature review of text summarisation.

**Dialogue Systems**

Compare and contrast the Conversational Analysis approach to dialogue with the Speech Act-based approach introduced in this module. Discuss which aspects of which approach are useful in building real dialogue processing systems.

**Applied Natural Language Processing**

Describe the concept of robust parsing. Why might robust parsing be useful? Describe a potential
application of robust parsing and using it as a reference, provide a literature review of robust parsing and associated natural language processing techniques.

**Programming Problems**

**Brill Part of Speech Tagger**

Write a Brill tagger which tags English text. You should provide a full evaluation of its performance. Brill taggers were described in Lecture 2. There's also a link to a technical paper by Eric Brill off the NLP and Applications webpage for you to read. Sample tagged text from the British National Corpus is available from the NLPA webpage.

**Finite State Transducer for Two Level Morphology**

Write a FST which can be used to parse English text into three levels: lexical, intermediate and surface. The FST should be able to deal with at least two types of inflection (for example plural versus singular for nouns and present and past for verbs). (as an alternative you could implement the porter stemmer as a FST - this is actually a lot of work but it's quite repetitive)

**CKY Parser for Probabilistic Context Free Grammars**

Write a CKY paper which accepts a Probabilistic Context Free Grammar and parses English text. You can either write your own PCFG, adapt one from the web (please provide a citation!) or adapt the grammar given in Jurafsky and Martin. The parser should be evaluated using the metrics described in Lecture 5.