Final Review

Scott Martin (Linguistics 384)

For the Exam on Thursday, December 6th, 2007
11:30 AM – 1:18 PM in 270 Journalism

1 Topics to be covered

- Grammar Correction (plus minimum edit distance for spell checking)
- Machine Translation
- Social context of language technology use

2 Format of the exam

You will have the entire 1:48 should you need/want it, but it should be doable to complete it in around one hour.

1. Term matching: (see below)

2. Calculations/Short answers:
    - Minimum edit distance
    - Venn Diagrams
    - Text alignment
    - Bag of words method
    - Context-free grammars
    - Syntactic trees
    - Transformer systems
    - Direct transfer systems
    - Interlinguas
    - Social context of language technology use

3 Terms to know

3.1 Grammar checking/context-free grammars (CFGs)

- syntax
- linear order
- constituents
- syntactic tree
- category
- lexical category
- morphology
- function word
- open class
- closed class
- phrase structure rule
- recursion
- directly and indirectly recursive rules
- generative
- structurally ambiguous
- hierarchical
- infinite
- context-free grammar
- pushdown automaton
- stack
- parsing
- local syntactic errors
- long-distance syntactic errors

3.2 Machine Translation

- translation
- machine translation
- (lexical) ambiguity
- structural ambiguity
- source language
- target language
- hyponym
- hypernym
- synonym
- interlingua
- translation triangle
- dictionary
- lexicon
- lexical gap
- light verb
- idiom
- reversibility
- transformer architecture
- direct transfer
- transfer component
- comparative grammar
- robustness
- text alignment
- sentence alignment
- word alignment
- machine learning
- training data
- bag of words
- underlying representation
- intelligibilty
- accuracy
- error analysis
- test suite
- taboo words

3.3 Social Context of (Language) Technology Use

- Deskilling
- Upskilling
- Metasearches
- Dewey decimal system
- Artificial intelligence