CS474
Introduction to Natural Language Processing
Midterm
October 27, 2005

Name:

Netid:

Instructions: You have 1 hour and 15 minutes to complete this exam. The exam is a closed-book exam.

```
#    description                               score         max score


1    word sense disambiguation               _____     /       15

2    p-o-s tagging                           _____     /       15

3    HMMs                                     _____     /       15

4    transformation-based learning / WSD     _____     /       20

5    bottom-up chart parsing                 _____     /       20

6    top-down parsing                        _____     /       15


Total score:                                 _____    /      100
```

1 Word Sense Disambiguation (15 pts)

Given the WordNet entries below, apply Lesk's dictionary-based word sense disambiguation algorithm to the word **bass** given the context **bass playing maniac**. For full credit, the step-by-step calculations must be shown and each step very briefly described.

bass
bass (the lowest portion of the musical range)
bass, basso (an adult male singer with the lowest voice)
sea bass, bass (the lean flesh of a saltwater fish of the family Serranidae)
bass (the member with the lowest range of a family of musical instruments)

playing
playing (the act of playing a musical instrument)
playing (the action of taking part in a game or sport or other recreation)

maniac
lunatic, madman, maniac (an insane person)
maniac (a person who has an obsession with or excessive enthusiasm for something)

Answer

After stemming and stop-word removal of all for the glosses, count the number of overlapping terms of each gloss of **bass** with all glosses associated with the context words.

- bass (the lowest portion of the musical range): score of 1

- basso (an adult male singer with the lowest voice): score of 0

- sea bass, bass (the lean flesh of a saltwater fish of the family Serranidae): score of 0

- bass (the member with the lowest range of a family of musical instruments): score of 2

So return the last sense of **bass**!

Comment

Some people compared each sense of bass with only one sense of each context word at a time, instead of comparing each sense of bass with all senses of all context words simultaneouly. (Needless to say, even though you compare one sense from the context words at a time, if you sum up the overlapping words for all senses of contenxt words before disambiguating bass, that is also correct.) In this example, either way you will get the same conclusion, but a small point (just -1!) is substracted for not being completely right. Also, (-1) for not mentioning stemming, and (-1) for not mentioning the stop word removal.

2   Part-of-Speech Tagging (15 pts)

Please read both parts of the problem before answering.

1. (7 pts) Briefly describe one potential problem with the Penn Treebank part-of-speech tagset.

   ## Answer

   There can be many good answers. Following are two possible answers. See also p.298 of J&M.

   A. Some special words can have only one possible tag regardless of its function in the sentence.
   B. Some sentences can be ambiguous.

2. (8 pts) Explain why the issue you raise in part (1) is a problem via a concrete example. Be sure to make clear how the example maps on to the problem you specified in part (1), e.g. you might show the relevant part-of-speech tags along with the words of an example sentence.

   ## Answer

   A.   • I went to/TO the park.
        • I went to/TO drive.

   Penn Treebank uses tag 'TO' for all occurrences of 'to', although the first 'to' is as preposition, and the second 'to' is as infinitive marker.

   B.   • Spring/ADJ leaves/NN spring/VB.
        • Spring/NN leaves/VBZ spring/NN.

   ## Comment

   (-4) for argueing without giving a concrete example.

3 HMM's for Part-of-speech Tagging (15 pts)

In the part-of-speech tagging assignment, you used Hidden Markov Models (HMMs) and the Viterbi algorithm to compute the most probable sequence of tags $T = t_1, t_2, \ldots, t_n$ given the words, $W = w_1 w_2 \ldots w_n$, in a sentence.

1. (8 pts) Assuming a **trigram** part-of-speech tagging model, provide the equation that is maximized by the Viterbi algorithm. (Be sure to define all variables and use clear notation.)

   ## Answer

   $$P(t_1)P(t_2|t_1) \prod_{i=3}^{n} P(t_i|t_{i-2}t_{t-1}) \prod_{i=1}^{n} P(w_i|t_i)$$

   ## Comment

   (-4) for giving only subproblem score functions for dynamic programming, but not the objective function as a whole. (-1) for not taking care of t=1 and t=2 cases carefully, or other minor notational problems.

2. (7 pts) Using Maximum Likelihood Estimation (and no smoothing), show (preferably via equations) how each term in the above trigram tagging model (from part 1) would be estimated from a training corpus.

   ## Answer

   $$P(t_i|t_{i-2}t_{i-1}) = \frac{C(t_{i-2}t_{i-1}t_i)}{C(t_{i-2}t_{i-1})}$$
   $$P(w_i|t_i) = \frac{C(w_i, t_i)}{C(t_i)}$$

4 Transformation-Based Learning (20 pts)

Think about how to best apply Brill's Transformation-Based Learning (TBL) algorithm (which we studied in the context of part-of-speech tagging) to the Senseval lexical tagging task for word sense disambiguation. (This is just the task you handled in the word sense disambiguation (WSD) assignment.)

1. (5 pts) What would be a reasonable **initial state tagger** for a TBL solution to the task?

   ## Answer

   Initialization with the most frequent sense for each word.

2. (5 pts) What **scoring function** should the transformation-based learner use for this task?

   ## Answer

   The score of each rule should capture the contribution of the rule when it is applied after the transformation by existing(thus far selected) set of rules. That is, (# of *new* correct instances - # of *new* incorrect instances).

   ## Comment

   (-2) for answers that didn't clearly point out the counting of correct/incorrect instances should be the delta(increment) with respect to the performance after transformation by the existing set of rules.

3. (10 pts) Design a set of tranformation **templates** for the task.

   ## Answer

   Ideal answers should provide templates that exploit informaiton based on words collocations and co-occurrences.

   (1) if w_i occurs within 10 words distance from w_0, change the sense of w_i to s_j.
   (2) if w_i occurs immediately after/before w_0, change the sense of w_i to s_j.
   (3) if the POS of w_i is p_l, change the sense of w_i to s_j.

# 5   Bottom-up Chart Parsing (20 pts)

Given the grammar and lexicon below, show the **final chart** for the following sentence after applying the bottom-up chart parser from class:

Run the Detroit marathon

Remember that the final chart contains all edges added during the parsing process. You may use either the notation from class (i.e. nodes/links) or the notation from the book to depict the chart.

S → NP VP	Det → the
S → VP	Noun → run | marathon
NP → Det NP	Verb → run
NP → Proper-Noun Noun	Proper-Noun → Detroit
VP → Verb NP	

Answer
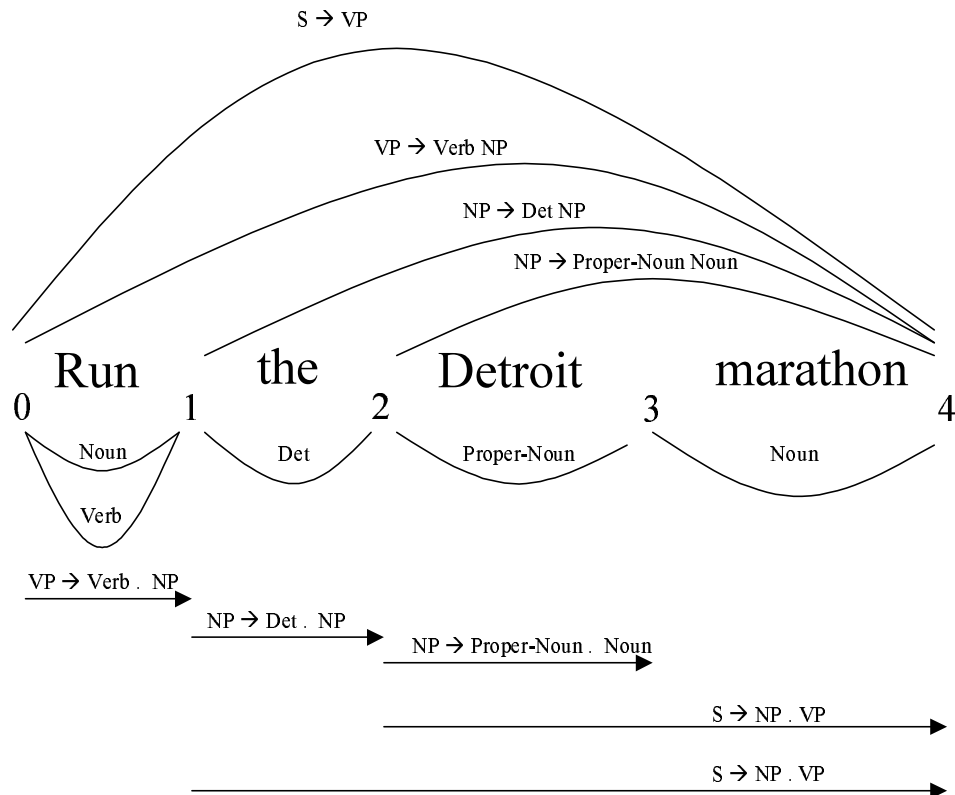
Comment

(-12) for giving a parse tree, and (-1 ˜ -4) for omitting various subsets of arcs.

6   Top-Down Parsing (15 pts)

1. Given the grammar and lexicon below (which is the same as that of question 5), show one possible **top-down derivation** for the sentence:

   Run the Detroit marathon

S → NP VP	Det → the
S → VP	Noun → run — marathon
NP → Det NP	Verb → run
NP → Proper-Noun Noun	Proper-Noun → Detroit
VP → Verb NP	

   ## Answer

   S → VP
   S → Verb NP
   S → Verb Det NP
   S → run Det NP
   S → run the Proper-Noun Noun
   S → run the Detroit Noun
   S → run the Detroit marathon

   ## Comment

   The questionn doesn't specify what parsing algorithm to use, so the order of application of the grammar rules can vary. (-4) for just giving a parse tree, because a parse tree cannot show whether the derivation is performed top-down or bottom-up. However, if you showed derivation with parse trees in a step by step manner, that is okey. Also, (-2) for showing the 'states transitions' of parsing, instead of 'derivation'.

2. The Earley algorithm introduces top-down predictions into the chart parsing algorithm. What top-down edge(s), if any, would be added to the chart at position 0 for the grammar and sentence above?

   ## Answer

   S → . NP VP
   S → . VP
   NP → . Det NP
   NP → . Proper-Noun Noun
   VP → . Verb NP