CSCI 5832 Spring 2015 Exam 1

Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

On my honor, as a University of Colorado at Boulder student, I have neither given nor received unauthorized assistance on this work. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **(5 points) True/False:** There are too many sunny days in Boulder. True or False.
2. **(5 points) True/False:** Any language described by a regular expression can also be captured by a deterministic finite state automaton. True.
3. **(5 points) Multiple choice:** Which of the following regular expressions describes the set of all lowercase alphabetic strings beginning with a ***b*** and containing at least one ***a***?
	1. /b[a-z]\*a+/
	2. /^b.\*a+[a-z]\*/
	3. /b[b-z]\*a[a-z]\*/ This one.
	4. /b([b-z]\*a[a-z]\*)\*/
4. **(10 points)** Describe in a sentence or two the language accepted by this FSA.



Any string that starts with ss or m and ends with one or more a’s.

1. **(5 points)** Add something to the FSA in problem 3 that would make it non-deterministic. Just draw on the figure above. Don’t worry if your amendment changes the language accepted by the machine. Many correct answers, such as an extra arc from q0 to q2 labeled with an s.
2. **(10 points)** How are the boldfaced words in the sentences below related to the word *elect?* If they are related in different ways, explain that difference.
	1. Hickenlooper was **elected** governor in 2012.
	2. The **election** had a high voter turn-out.

*Elected* adds the inflectional morpheme *–ed* to the stem elect, changing its tense. *Election* adds a derivational morpheme, changing the verb *elect* to a noun. They are different processes. Inflections do not change the word class, and the new word meaning is transparently related to the original. Derivational morphemes often change the word class and the new meaning is not always predictable or transparently related to the original.

1. **(5 points)** True/False: The minimum edit distance algorithm is an example of dynamic programming because it computes the cost of every possible path. False.
2. Think about the task of part of speech tagging and the use of Bayes rule to find the most probable tag sequence for a given sequence of words.
	1. **(5 points)** What intuition about language is represented by the prior **p (ti|ti-1)** ? Give an example of **(ti|ti-1)** with a high probability in English.

That some pos tags are more likely to follow a particular tag than others because there are regular grammatical patterns in language. E.g., the probability of a noun following a determiner is high.

* 1. **(5 points)** What intuition about language is represented by the likelihood

**p (wi|ti)** ? Give an example of **(wi|ti)** with a high probability in English.

That a given part of speech is more likely to be instantiated with certain words than others because of e.g., pragmatics, Zipf’s law, etc. One example: the 3rd person singular verb tag is much more likely to appear with the word *is* than with the word *wash.*

9. False. p(Observations|Model) is computed with the forward algorithm

10. False. While there are S^O number of paths, the Viterbi algorithm, using dynamic programming, does not consider all of them. The actual number of operations is in the order of O\*S^2 (considers all possible single state transitions for each observation step), although that was not required knowledge to correctly answer the question.

11

a) the expected counts of i to j transitions over expected counts of i to any next state (slide 36, 38 on the Thursday lecture)

b) Maximization step in the Expectation-Maximization algorithm (slide 39 in the Thursday lecture), which forward-backward algorithm is an example of. The question was a bit vague, so partial credit was awarded if the answer resembled slide 40 on the Thursday lecture. However, there is no separate "backward" step. Forward-backward refers to computing the forward probability (used to re-estimate transition probability a\_ij) and backward probability (used to re-estimate the observation probability).