START OF QUIZ Student ID: 49919301,Keigan,Jonathan

Question 1

Topic: Lecture 2
Source: Lecture 2

Describe the purpose of linkage in hierarchical clustering. (1)

Question 2

Topic: Lecture 1
Source: Lecture 1

Suppose we are filling the table for the Levenshtein distance algorithm. We are in cell (x, y). The values of cell (x-1, y-1), (x-1, y), and (x, y-1) are 3, 4, and 3, respectively. What is the value we will put in cell (x, y), given that the letters are equal? (1)

Question 3

Topic: Lecture 4
Source: Lecture 4

How is it that EM can arrive at a good solution, even if we have a random initialization of parameters? (1)

Question 4

Topic: Lecture 3
Source: Lecture 3

Describe the noisy channel model, and how it can be used to represent POS-tagging. (1)

Question 5

Topic: Lecture 1
Source: Lecture 1

When is dynamic programming more efficient than brute force programming? (ie, what assumptions do we make about a problem when we use dynamic programming?) (1)

Question 6

Topic: Lecture 4
Source: Lecture 4

Imagine that we are doing OCR (optical character recognition; ie, the translation of handwritten text into digital text) instead of POS tagging. Do you think we could use an HMM? If so, what would the states, transitions, and emissions be? If not, describe why it's an inappropriate tool for the task. (2)

Question 7

Topic: Lecture 3
Source: Lecture 3

Imagine that we have a trigram model that encounters a trigram where none of the tokens are in the vocabulary. How do you think that might impact our probability calculation for the sentence? How might we go about finding a solution? (2)

Question 8

Topic: Lecture 2
Source: Lecture 2

Imagine we were using k-means to cluster misspellings around their correct spellings. How many clusters would we need, and what would be a good distance function? Explain. (2)

Question 9

Topic: Long
Source: Lecture 3

In class, we built a collocation matrix for a bigram language model. Modify the function so that it can handle a trigram language model and implements “add-alpha” smoothing, instead of “add-one” smoothing. (3)

END OF QUIZ