START OF QUIZ
Student ID: 42135814,Lopez Gonzalez,Nico

Question 1

Topic: Lecture 2
Source: Lecture 2

Do you think we could build a POS tagger with an FST, where one side of the tape is POS tags, and the other side of the tape is words? Explain why / why not. (1)

Question 2

If we were to try to use an HMM for segmentation, describe what the transition and emission probabilities would be. (1)

Question 3

Topic: Lecture 1
Source: Lecture 1

Describe allomorphy, with an example we did not cover in class. (1)

Question 4

Topic: Lecture 4
Source: Lecture 4

Transition-based segmentation is very similar to the SR parser we saw last block, except it uses 2 FIFO structures, and doesn't require a stack. What is different about segmentation so that it doesn't require a stack? (1)

Question 5

What is the role of the partition function $(Z(x))$ in CRFs? Why is it necessary? (1)

Question 6

Topic: Lecture 1
Source: Lecture 1

Vowel harmony is a process by which vowels in affixes must match some of the properties of the vowels in the root. For example, in Turkish, "houses" is "evler", while "schools" is "okullar", where the plural suffix must have a front or back vowel, matching the root ("ev" and "okul"). Given the ML models you've seen so far, give a specific example of a model that you think can learn this process, and explain why it's well suited to the task. (2)

Question 7

Topic: Lecture 3
Source: Lecture 3

In the lab, you compared BPE with a more linguistically-motivated segmentation scheme. Intrinsically, the supervised method performs much better, but typically, BPE and its cousins work much better down-stream. Why do you think that is, taking into account the differences between the two methods? (2)

Question 8

Topic: Lecture 2
Source: Lecture 2

As a thought experiment, how might we build a calculator using an FST? Imagine that the FST reads input on one side of the tape, and generates operations (that are carried out by an algorithm) on the output side. (2)

Question 9

Topic: Long
Source: Lecture 1

In English (and other stress-timed languages, such as German, Russian, Arabic, Greek, Hindi, Thai, etc.), stress tends to be strongest on the root of the word, and is softer along affixes (and in English, on periphrastic necessities like auxiliary verbs - try it!). What implications might this have on an ASR system, do you think they are a significant issue, and can you envision any way of moderating them with morphological knowledge? (3)

END OF QUIZ