START OF QUIZ
Student ID:
50656347,Zhang,Lisa

Question 1

Topic: Lecture 4
Source: Lecture 4

Literate Chinese speakers have no difficulty parsing words in Chinese text (outside normal ambiguities). What advantages do you think they have that our algorithms are lacking? (1)

Question 2

What kind of tasks are CRFs appropriate for (2 requirements)? Don't list tasks, but rather the general class of tasks. (1)

Question 3

Knowing what you know about parsing, describe how compounding could be considered syntax, instead of morphology. In other words, how might we parse compounds? (1)

Question 4

Topic: Lecture 3
Source: Lecture 3

Why is differential entropy a good measure for establishing morpheme boundaries? (1)

Question 5

Topic: Lecture 3
Source: Lecture 3

If we were to try to use an HMM for segmentation, describe what the transition and emission probabilities would be. (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 2
Source: Lecture 2

In some ways, Statistical Machine Translation (SMT) was similar to an FST modified by a re-ordering model (ie, each input word had a corresponding output translation, and then the words were re-ordered to fit a language model). These models have been supplanted by NMT. What shortcomings of FSTs do you think put a ceiling on SMT performance? (2)

Question 8

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

There is an alternative to BPE that randomly "forgets" to merge together certain subword sequences when it is creating its vocabulary (for example, "forget" will occasionally be represented as "for-get", occasionally as "forget", occasionally as "for-g-et", etc. What impacts do you think this might have on the vocabulary and model performance? Secondly, do you think there is a different impact between forgetting early iteration, mid iteration, and late iteration merges? (3)

END OF QUIZ