

**START OF QUIZ**

**Student ID:**

**97713317, Song, Shawn**

## Question 1

Topic: Lecture 1

Source: Lecture 1

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

## Question 2

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 3

Topic: Lecture 2

Source: Lecture 2

Do you think that FSTs can work with templatic morphology? Explain. (1)

## Question 4

Topic: Lecture 4

Source: Lecture 4

Do you think that we could do Chinese Word Segmentation in a bottom-up way like we do with BPE? Why might this work (or not)? (1)

## Question 5

Topic: Lecture 3

Source: Lecture 3

Describe, with an example (not used in class) the difference between surface and canonical segmentation. (1)

## Question 6

Topic: Lecture 4

Source: Lecture 4

Garden path sentences are sentences that start with one parse, but need to be reparsed in the middle of the sentence (“The old man the boats.” - ‘old’ changes from an adjective to a noun, and ‘man’ from a noun to a verb). A bad Chinese word segmentation could result in the same need to re-parse our segmentation after encountering a new word. Of the methods we looked at, which do you think is the most likely to be able to “correct” a segmentation? Explain. (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

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 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**