

**START OF QUIZ**

**Student ID:**

**37157856, Wang, Lusha**

## Question 1

Topic: Lecture 3

Source: Lecture 3

In class, we saw that the entropy of a fair coin toss is 1 (bit), but that was because we were using a binary logarithm. Similarly, the entropy of an 8-sided fair die is 1, if we use an octal logarithm. What is the entropy of an 8-sided die using a binary logarithm? Either show your work or explain the relation. (1)

## Question 2

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 3

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 4

Topic: Lecture 1

Source: Lecture 1

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

## Question 5

Topic: Lecture 1

Source: Lecture 1

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 6

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 7

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 8

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 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**