

**START OF QUIZ**

**Student ID:**

**70152863,Khan,Muhammad  
Mujtaba**

## Question 1

Topic: Lecture 2

Source: Lecture 2

Umlaut is a morpho-phonological process that moves a vowel forward in the mouth under certain morphological processes (for example, Hund+PL -> Hünde in German). What might the re-write rule for this example look like? (1)

## Question 2

Topic: Lecture 2

Source: Lecture 2

Do you think that FSTs can work with vowel harmony? Explain. (1)

### Question 3

Topic: Lecture 1

Source: Lecture 1

English is often described as an “analytic language with some fusional properties”. Describe what that means, with an example. (1)

## Question 4

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 5

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 6

Topic: Lecture 4

Source: Lecture 4

Imagine that we had a language like Chinese that doesn't use spaces for word segmentation, but is considerably more morphologically complex than Chinese. Do you think that it would be easier or harder to segment? Give an explanation (and include any assumptions). (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 1  
Source: Lecture 1

Suppletion is a process by which morphological patterns (called paradigms) merge to form a mixed paradigm. For example, the past tense of “to go” comes from an older verb, “wendan - to turn”. Describe how syncretic paradigms might impact a machine learning model, and how we can learn to model them accurately. (2)

## Question 9

Topic: Long

Source: Lecture 3

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**