# START OF QUIZ
# Student ID:
# 18643544,Zhang,Cindy

## Question 1

Topic: Lecture 8
Source: Lecture 8

Why do you think that LLMs learn syntax earlier in the model than morphology? (1)

## Question 2

Topic: Lecture 8
Source: Lecture 8

Where do you think pragmatic learning (ie, intent) might fall within the layers of an LLM? Explain briefly. How might we test for it? (1)

## Question 3

Topic: Lecture 7
Source: Lecture 7

What benefits would evaluating an inflection model on nonce words have? Are there any disadvantages? (1)

## Question 4

Topic: Lecture 5
Source: Lecture 5

Why do we attach an embedding layer before passing information to the hidden layer(s)? (1)

## Question 5

Topic: Lecture 5
Source: Lecture 5

Why is POS information so important (whether via tagging or embedded information)? (1)

## Question 6

Topic: Lecture 6
Source: Lecture 6

Feature engineering can be incorporated into encoder-decoder models through the use of multiple encoders. If you could have any extra annotation for morphological analysis, and were able to pass each through a separate encoder, what types of features would you include? Do you see any potential problems with using this extra annotation? (2)

## Question 7

Topic: Lecture 6
Source: Lecture 6

We know that domain shift can have a significant impact on the quality of our models - despite POS tagging being an "easy" task, POS taggers fail spectacularly when we try to use them on different domains. Do you think the same would be true of (contextual) morphological analysers? What similarities and differences between POS and MSD led you to this conclusion? (2)

## Question 8

Topic: Lecture 7
Source: Lecture 7

What benefits might encoding MSDs with a second encoder have over a single encoder approach? Can you think of any disadvantages? (2)

## Question 9

Topic: Long
Source: Lecture 5

In class, we talked about how POS and morphological information is often latently encoded in word embeddings, but not in character embeddings. Let's think about subword embeddings, since most DL models are going to use subword representations. If a word is split, where do you think this information is encoded, and does it matter? Explain your reasoning. (3)

# END OF QUIZ