CS685 Quiz 4: *probe tasks*
Released 4/13, due 4/19 on Gradescope (please upload a PDF!)
Please answer both questions in 2-4 sentences each.

1. Let's say we want to design a probe task to measure whether BERT's token-level representations encode information about whether or not a particular token is a named entity. Why do we have to freeze BERT's representations during the training of our probe network?

Because we're interested on probing the encoder. Whatever it's Bert or Roberta. We don't want its weights to be updated during the training. We'll be only training the classification head.

2. Now let's say we want to probe whether or not BERT's [CLS] token has encoded the length of an input sentence. Explain how you would design a control task for this probe to address the effect of probe network complexity.