HW6

Olga Whelan (olgaw@uw.edu)
Mike Roylance (roylance@uw.edu)

Q1

In your note file, explain what the states and the transition and the emission probs in the output-hmm represent:

The states represent the part of speech tags that we have encountered. For each tag, we have one state. We also include the beginning of string (BOS).

The transition probability is the probability of one part of speech being followed by another. For example, if a noun followed a modal verb, we would have a transition of P(N | MV).

The emission probability is the probability of getting a particular word with a specific part of speech. For instance, if we had pronouns and we only found “he” and “she” with equal distribution, then we should have P(he | pronoun) = 0.5 and P(she | pronoun) = 0.5.

Q2

In your note file, explain what the states and the transition and the emission probs in the output-hmm represent:

The states represent the part of speech tags that we have encountered. For each tag, we have one state. We also include the beginning of string (BOS).

The transition probability is the probability of one part of speech being preceded by two other parts of speech. This probability is smoothed with the big formula from the slides.

The emission probability is the probability of getting a particular word with a specific part of speech. This is using the formula from the slides as well, where we calculate what the probability is from the unknown probability file.