CS224N: Natural Language Processing
Spring 2008
Weekly Quiz

Quiz 2

1: The beam search technique discussed in J&M, when deciding which sentences to expand at each step, compares scores between candidate translation sentences that have different numbers of words by:

- Normalizing the score by the length of the sentence
- Comparing scores on sub-sections of the sentences that have equal length
- Comparing only scores between sentences with the same number of words
- None of those above
- All of those above

2: Which of the following is true about EM?

- If you run it long enough, it eventually reaches a global optimum.
- Different initial condition may lead to different results.
- You have to have completely observed data to do EM.
- The data likelihood sometimes goes down after an iteration.
- None of those above

3: Which of the following is FALSE about IBM Models?

- Adding the fertility model makes it hard to train IBM Model 3 efficiently
- IBM Model 1, even though simple, takes a long time because it has to sum over all possible alignment structures of a sentence pair.
- IBM Model 2 can be efficiently computed even though it makes use of a distortion model
- With IBM Model 3, you need to initialize the model parameters with some good estimates instead of just random or uniform ones.
- None of the above

4: Which of the following is FALSE?

- The MT automatic evaluation metric BLEU is a weighted average of the number of N-gram overlaps with the human translations.
- Having multiple human reference translation makes BLEU more reliable
- Since BLEU is computing n-gram overlaps, a system can get a high BLEU score by translating every foreign sentence to one very common word (e.g., "the"), obtaining a high unigram precision.

- None of those above
- All of those above

5: We will examine $P_{continuation}(w)$ for the given, incomplete sentence:

"How much wood would a woodchuck chuck if a woodchuck could chuck"

What is $|\{w_{i-1} : C(w_{i-1}\ w_i)>0\}|$ for w_i="woodchuck"?

- 0
- 1
- 2
- 3

6: Which word is more likely to complete the sentence (follow the last "chuck") based on $P_{continuation}$?

- How
- wood
- would
- chuck

7: Now we will use the modified sentence below where all instances of "woodchuck" have been misspelled as "wood chuck":

"How much wood would a wood chuck chuck if a wood chuck could chuck"

Now what is $|\{w_{i-1} : C(w_{i-1}\ w_i)>0\}|$ for w_i="chuck"?

- 0
- 1
- 2
- 3

8: Now which word is more likely to extend this sentence based on $P_{continuation}$?

- How
- wood
- would
- chuck

9: We will explore a simple example for using EM to train alignment models. Say we have the following vocabularies for the two languages: {A,B,C} and {w,x,y,z}; and the following three pairs of sentences:

`A C`	`A B`	`B B C`
`x y`	`x w`	`w z y`

We use a simplified version of Model 1 in which we ignore the NULL word and alignments where there are spurious or zero-fertility words. We initialize with

uniform weights for all translation probabilities, e.g., t(w|A) = t(x|A) = t(y|A), etc., and t(w|A) = t(w|B) = t(w|C). Thus every alignment is equiprobable, and there are 2 distinct alignments for each of the first two sentence pairs and 6 distinct alignments for the third sentence pair. In the first iteration of EM, we collect fractional counts for each alignment. What fractional count do we collect for (A|w) (what is tcount(A|w))?

- ◯ 1/4
- ◯ 1/3
- ◯ 1/2
- ◯ 1
- ◯ None of the above

10: What fractional count do we collect for (w|B) (what is tcount(w|B))?

- ◯ 1/2+1/6
- ◯ 1/2+1/3
- ◯ 1/2+1/2
- ◯ 1/2+2/3