1. The HMM graph is roughly represented below:

start -> cow : 1

cow -> cow : 0.5, cow -> duck : 0.3, cow -> end : 0.2

duck -> duck : 0.5, duck -> cow : 0.3, duck -> end : 0.2

a) The viterbi matrix:

|  |  |  |  |
| --- | --- | --- | --- |
| State/Observation | moo | hello | quack |
| q0 |  |  |  |
| q1 (Cow) | 0.9 | max(0.9\*0.5\*0.1, 0) = 0.045 | max(0.045\*0.5\*0, 0.108\*0.3\*0) = 0 |
| q2 (Duck) | 0 | max(0.9\*0.3\*0.4, 0) = 0.108 | max(0.045\*0.3\*0.6, 0.108\*0.5\*0.6) = 0.0324 |
| q3 |  |  | max(0, 0.0324\*0.2) = 0.00648 |

The backpointer matrix:

|  |  |  |  |
| --- | --- | --- | --- |
| State/Observation | moo | hello | quack |
| q0 |  |  |  |
| q1 | 0 | 1 | 2 |
| q2 | 0 | 1 | 2 |
| q3 |  |  | 2 |

So the decoded sequence is moo/Cow hello/Duck quack/Duck, and the probability is 0.00648.

b) Yes there is one, and the sequence is moo/Cow hello/Cow quack/Duck with its probability 0.00162. So the total probability of emitting this sentence is 0.0081.

2.

Correctly tagged sentence is “I like singing.”, and incorrectly tagged sentence is “I like milk.”. Their outputs are listed below:

---------

Sentence: I like singing.

Annotate I == <constit cat="PRP">

Annotate like == <constit cat="VBP">

Annotate singing == <constit cat="VBG">

Annotate . == <constit cat=".">

---------

Sentence: I like milk.

Annotate I == <constit cat="PRP">

Annotate like == <constit cat="IN">

Annotate milk == <constit cat="NN">

Annotate . == <constit cat=".">

The only incorrectly tagged word in the second sentence is “like”. In order to demonstrate why it is incorrectly tagged by the tagger using the training corpus, I am going to show the only different part of the probability product starting from state transition from the tag of “I” to the tag of “like” and ending at the state transition from the tag of “like” to the tag of “milk” between the incorrect tag sequence “PRP IN NN” and the correct one “PRP VBP NN” below:

PRP IN NN: (735 / 20819) \* (475 / 118291) \* (12815 / 118291) = 1.535805 \* 10^-5

The parenthesized terms are the transitional probability from PRP to IN, the emission probability of “like” of IN, and the transitional probability from IN to NN, respectively.

PRP VBP NN: (3659 / 20819) \* (25 / 14955) \* (447 / 14955) = 0.878167 \* 10^-5

The parenthesized terms are the transitional probability from PRP to VBP, the emission probability of “like” of VBP, and the transitional probability from VBP to NN, respectively.

We can see that the correct tagged sequence has smaller product than the incorrect one.