Homework #2A

In Rabiner's tutorial paper on HMMs, Rabiner desribes a HMM of *n* hidden states $\{S_1, S_2, \ldots, S_n\}$ and *m* possible observations $\{v_1, v_2, \ldots, v_m\}$ in terms of **λ =(π , A , B)** where (i) π is an *n*-by-1 vector encoding the initial-state probabilities, (ii) A is an *n*-by-*n* matrix encoding the state transition probabilities, and (iii) B is an *n*-by-*m* matrix encoding the observation probabilities. **For Homework #2A, λ =(π , A , B) is shown in the answers to Question #1 and Question #2.**

Question #1:

Answers: (5 points, 1 point for π and 4 points for A)

Vector *π* = <1, .0, 0, 0, 0>

where columns are ordered corresponding to the five states S_1=*I*, S_2=*a*, S_3=*b*, S_4=c, S_5=*F*.

Matrix A =

0	.57	.29	.14	0
0	.2	.46	.23	.11
0	0	.2	.53	.27
0	0	0	.2	.8
0	0	0	0	1

where both rows and columns are ordered corresponding to the six states

S_1=*I*, S_2=*a*, S_3=*b*, S_4=c, S_5=*F*.

Question #2:

Answer: (5 points for B**)**

Matrix B =

1	0	0	0	0	0
0	.9	.04	.02	.04	0
0	.04	.9	.04	.02	0
0	.02	.04	.9	.04	0
0	0	0	0	0	1

where rows are ordered corresponding to the states S_1=*I*, S_2=*a*, S_3=*b*, S_4=c, S_5=*F*

while columns are ordered corresponding to the observation symbols

v_1=*ReadyToType*, v_2=*a*, v_3=*b*, v_4=c, v_5=d, v_6=*EndOfWord.*

Question #3

Answer: (2 points)

(1.0*.57*.46*.2*.27) * (1.0*.04*.9*.02*1.0) = 0.00001019433.

Explanations:

Given the HMM λ =(π , A , B) shown in the answers to Question #1 and Question #2,

what is the probability Pr(O, Q | λ) of going through

a given state sequence Q: I ➔a➔b➔b➔F and

seeing a sequence of observations O: ReadyToType➔b➔b ➔d➔EndOfWord ?

Steps:

- According to equation 14,
  for Q= I ➔a➔b➔b➔F in our case,
  the probability of the state sequence Q given the HMM λ =(π , A , B) is
  $Pr(Q \mid \lambda) = \pi_I * a_{Ia} * a_{ab} * a_{bb} * a_{bf}$ = 1.0*.57*.46*.2*.27.
- According to equation 13a and 13b,
  for Q= I ➔a➔b➔b➔F and
  O = ReadyToType➔b➔b ➔d➔EndOfWord in our case,
  the probability of the onservation sequence O
  given the state sequence Q and HMM λ =(π , A , B) is
  $Pr(O|Q, \lambda) = b_{I\ ReadyToType} * b_{a\ b} * b_{b\ b} * b_{b\ d} * b_{F\ EndOfWord}$ = 1.0*.04*.9*.02*1.0.
- According to equation 15, the probability Pr(O, Q | λ) equals Pr(Q| λ) * Pr(O|Q, λ),
  for Q= I ➔a➔b➔b➔F and
  O = ReadyToType➔b➔b ➔d➔EndOfWord in our case,
- (1.0*.57*.46*.2*.27) * (1.0*.04*.9*.02*1.0) = 0.00001019433

Question #4

Answer: 3 points (0.3 point for each of the 10 possible state transition sequences)

Theoretically there are 3*3*3=27 possible state transition sequences in the form of I -> ? -> ? -> ? -> F. However, a number of them have transitions from b to a, c to a, or c to b and end in probability 0. Excluding these, the following are the 10 possible state transition sequences Q, i.e. those with none zero probability.

I -> a -> a -> a -> F

I -> a -> a -> b -> F

I -> a -> a -> c -> F

I -> a -> b -> b -> F

I -> a -> b -> c -> F

I -> a -> c -> c -> F

I -> b -> b -> b -> F

I -> b -> b -> c -> F

I -> b -> c -> c -> F

I -> c -> c -> c -> F

Question #5:

Answer: (5 points if the final answer is correct. Or give it 0.5 point for each correctly calculated

Pr(O, Q | λ) for each of the 10 possible state transition sequences Q if they are shown.)

0.00105680007 (i.e. about 0.001057)

Explanations:

In this case, the observations sequence O = *ReadyToType*➔*b*➔*b*➔*d* ➔*EndOfWord*.

The probability of seeing *bbd* as the result of the person trying to type the word *abc* is Pr(O| λ), where O is the observations sequence O = *ReadyToType*➔*b*➔*b*➔*d* ➔*EndOfWord* and λ =(π , A , B) is the HMM encoding the spelling model for the word *abc* and keyboard model.

According to equation 16, Pr(O| λ) is simply the sum of all Pr(O, Q | λ) for all possible state sequences Q enumerated in Question #4.

Therefore we can repeat the approach in Question #3

and

for O = *ReadyToType➔b➔b➔d ➔EndOfWord* and
for each Q enumerated in Question #4

determine probability Pr(O, Q | λ) as shown in the following:

I -> a -> a -> a -> F	(1*.57*.2*.2*.11)*(1*.04*.04*.04*1) = 0.000000160512	0.000000161
I -> a -> a -> b -> F	(1*.57*.2*.46*.27)*(1*.04*.04*.02*1) = 0.0000004530816	0.000000453
I -> a -> a -> c -> F	(1*.57*.2*.23*.8)*(1*.04*.04*.04*1) = 0.00000134246	0.000001342
I -> a -> b -> b -> F	(1*.57*.46*.2*.27)*(1*.04*.9*.02*1) = 0.00001019433	0.000010194
I -> a -> b -> c -> F	(1*.57*.46*.53*.8)*(1*.04*.9*.04*1) = 0.00016008883	0.000160089
I -> a -> c -> c -> F	(1*.57*.23*.2*.8)*(1*.04*.04*.04*1) = 0.00000134246	0.000001342
I -> b -> b -> b -> F	(1*.29*.2*.2*.27)*(1*.9*.9*.02*1) = 0.0000507384	0.000050738
I -> b -> b -> c -> F	(1*.29*.2*.53*.8)*(1*.9*.9*.04*1) = 0.0007967808	0.000796781
I -> b -> c -> c -> F	(1*.29*.53*.2*.8)*(1*.9*.04*.04*1) = 0.00003541248	0.000035412
I -> c -> c -> c -> F	(1*.14*.2*.2*.8)*(1*.04*.04*.04*1) = 0.000000286727	0.000000287

According to equation 14,

the probability of seeing *bbd* as the result of the person trying to type the word *abc* ,

i.e. Pr(O| λ) is equal to the sum of the probabilities above:

0.000000161+

0.000000453+

0.000001342+

0.000010194+

0.000160089+

0.000001342+

0.000050738+

0.000796781+

0.000035412+

0.000000287+

=

0.00105680007 (i.e. about 0.001057)