CS 271 Spring 2020
Assignment 0

Student ID: 014498887
Name: Yash Sahasrabuddhe

Q.1. a)

Hidden State X	**P(O,λ)**	**Normalized probability**
HHH	0	0
HHC	0	0
HCH	0	0
HCC	0	0
CHC	1 * 0.2 * 0.4 * 0.1 * 0.3 * 0.1 = 0.00024	0.009
CCH	1* 0.2 * 0.6 * 0.7 * 0.4 * 0.5 = 0.01680	0.6765
CHH	1 * 0.2 * 0.4 * 0.1 * 0.7 * 0.5 = 0.0028	0.1125
CCC	1 * 0.2 * 0.6 * 0.7 * 0.6 * 0.1 = 0.00504	0.202
Sum	**0.02488**	**1**

Table 1

Using the direct formula to calculate the probability of the hidden state sequences, we use the formula $\pi_{x0}b_{x0}(O_0)a_{x0,x1}b_{x1}(O_1)a_{x1,x2}\, b_{x2}(O_2)$. Here, π_{x0} is the initial state's probability followed by $b_{x0}(O_0)$ which is the probability of the observation occurring in state b_{x0} and $a_{x0,x1}$ being the probability of transitioning from state 0 to state 1.

Similarly, doing this calculation for all the observations is used to calculate the values of each **P(O,λ)** for each possible state sequence for that observation.

b)

$\alpha(i)$	$P(O,\lambda)$
$\alpha_0(0)$	0 * 0.4 = 0
$\alpha_0(1)$	1 * 0.2 = 0.2
$\alpha_1(0)$	(0 * 0.7 + 0.2 * 0.4) * 0.1 = 0.008
$\alpha_1(1)$	(0 * 0.3 + 0.2 * 0.6) * 0.7 = 0.084
$\alpha_2(0)$	(0.008 * 0.7 + 0.084 * 0.4) * 0.5 = 0.0196
$\alpha_2(1)$	(0.008 * 0.3 + 0.084 * 0.6) * 0.1 = 0.00528
	Total = 0.11688

For the initial state, we use the formula $\boldsymbol{\alpha_0(i) = \pi_i b_i(O_0)}$, for initialization step and then use the following recurrence equation to calculate the next values-

$$\alpha_t(i) = \left(\sum_{j=0}^{N-1} \alpha_{t-1}(j)a_{ji}\right) b_i(\mathcal{O}_t)$$

where-
$\boldsymbol{\alpha_{t-1}}$ is the previous α value we calculated, $\boldsymbol{a_{ji}}$ is the state transition probability and $b_i\,(O_t)$, Using these formulas and the probability matrices, we can calculate the **P(O,λ)** using the α-pass algorithm.

c) The work factor in the first computation is very high as there is no recursion and we are calculating P(O,λ) directly. Hence, it's work factor is $2TN_T$.

On the other hand, the α-pass or forward algorithm is calculating P(O,λ) recursively which reduces the number of computation massively which only requires only about N_2T multiplications worth of work factor.

Q.2. a) Referring to the table 1 above, for dynamic programming sense, we will just pick the state sequence which has the best probability i.e. the highest probability. In this case, the best hidden state sequence in DP sense is **CCH** as it has the highest probability.

b) For HMM's best hidden state sequence, we will find the best probability for each of the position. First, we will see the best probability at the first position. For this we will see the probability of the two sets of sequences which have either 'H' at the first position or 'C' at the first position. We see that, the sum of probabilities of the sequences starting with 'C' is greater. Hence, 'C' is the most suitable for the first position.

Similarly, we will check all the sequence's probabilities for each of the position and sum it up and see which one is greater. Finally, we will see such a matrix,

State (X)	Position 0	Position 1	Position 2
P(H)	0	0.1215	0.789
P(C)	1	0.8785	0.211

Q.3. a) Using the Python code attached in the mail (DComputation.py) i.e. the direct computation calculation of the probabilities, the following probabilities were calculated for each observation sequence-

Observation Sequence	Probability
[0, 0, 0, 0]	0.02840236
Observation Sequence	Probability
[0, 0, 0, 1]	0.01835428
Observation Sequence	Probability
[0, 0, 0, 2]	0.01744736
Observation Sequence	Probability
[0, 0, 1, 0]	0.0146524
Observation Sequence	Probability
[0, 0, 1, 1]	0.0131824
Observation Sequence	Probability
[0, 0, 1, 2]	0.0138572
Observation Sequence	Probability
[0, 0, 2, 0]	0.01260644
Observation Sequence	Probability
[0, 0, 2, 1]	0.01300292
Observation Sequence	Probability
[0, 0, 2, 2]	0.01409464
Observation Sequence	Probability
[0, 1, 0, 0]	0.0147616
Observation Sequence	Probability
[0, 1, 0, 1]	0.0099688
Observation Sequence	Probability
[0, 1, 0, 2]	0.0096296
Observation Sequence	Probability
[0, 1, 1, 0]	0.0102496
Observation Sequence	Probability
[0, 1, 1, 1]	0.0101728
Observation Sequence	Probability
[0, 1, 1, 2]	0.0109376
Observation Sequence	Probability
[0, 1, 2, 0]	0.0099968
Observation Sequence	Probability
[0, 1, 2, 1]	0.0110024
Observation Sequence	Probability

[0, 1, 2, 2]	0.0120808
Observation Sequence	Probability
[0, 2, 0, 0]	0.01274924
Observation Sequence	Probability
[0, 2, 0, 1]	0.00880052
Observation Sequence	Probability
[0, 2, 0, 2]	0.00856624
Observation Sequence	Probability
[0, 2, 1, 0]	0.010022
Observation Sequence	Probability
[0, 2, 1, 1]	0.0102608
Observation Sequence	Probability
[0, 2, 1, 2]	0.0111052
Observation Sequence	Probability
[0, 2, 2, 0]	0.01016356
Observation Sequence	Probability
[0, 2, 2, 1]	0.01138708
Observation Sequence	Probability
[0, 2, 2, 2]	0.01254536
Observation Sequence	Probability
[1, 0, 0, 0]	0.0194936
Observation Sequence	Probability
[1, 0, 0, 1]	0.0126728
Observation Sequence	Probability
[1, 0, 0, 2]	0.0120736
Observation Sequence	Probability
[1, 0, 1, 0]	0.01052
Observation Sequence	Probability
[1, 0, 1, 1]	0.009632
Observation Sequence	Probability
[1, 0, 1, 2]	0.010168
Observation Sequence	Probability
[1, 0, 2, 0]	0.0092584
Observation Sequence	Probability
[1, 0, 2, 1]	0.0096712
Observation Sequence	Probability
[1, 0, 2, 2]	0.0105104
Observation Sequence	Probability
[1, 1, 0, 0]	0.0142016
Observation Sequence	Probability
[1, 1, 0, 1]	0.0097568
Observation Sequence	Probability
[1, 1, 0, 2]	0.0094816
Observation Sequence	Probability
[1, 1, 1, 0]	0.01088
Observation Sequence	Probability

[1, 1, 1, 1]	0.011072
Observation Sequence	Probability
[1, 1, 1, 2]	0.011968
Observation Sequence	Probability
[1, 1, 2, 0]	0.0109504
Observation Sequence	Probability
[1, 1, 2, 1]	0.0122272
Observation Sequence	Probability
[1, 1, 2, 2]	0.0134624
Observation Sequence	Probability
[1, 2, 0, 0]	0.0140728
Observation Sequence	Probability
[1, 2, 0, 1]	0.0098344
Observation Sequence	Probability
[1, 2, 0, 2]	0.0096128
Observation Sequence	Probability
[1, 2, 1, 0]	0.0118
Observation Sequence	Probability
[1, 2, 1, 1]	0.012256
Observation Sequence	Probability
[1, 2, 1, 2]	0.013304
Observation Sequence	Probability
[1, 2, 2, 0]	0.0121832
Observation Sequence	Probability
[1, 2, 2, 1]	0.0137576
Observation Sequence	Probability
[1, 2, 2, 2]	0.0151792
Observation Sequence	Probability
[2, 0, 0, 0]	0.01893724
Observation Sequence	Probability
[2, 0, 0, 1]	0.01233652
Observation Sequence	Probability
[2, 0, 0, 2]	0.01176224
Observation Sequence	Probability
[2, 0, 1, 0]	0.0103756
Observation Sequence	Probability
[2, 0, 1, 1]	0.0095536
Observation Sequence	Probability
[2, 0, 1, 2]	0.0100988
Observation Sequence	Probability
[2, 0, 2, 0]	0.00919796
Observation Sequence	Probability
[2, 0, 2, 1]	0.00964628
Observation Sequence	Probability
[2, 0, 2, 2]	0.01049176
Observation Sequence	Probability

[2, 1, 0, 0]	0.0151648
Observation Sequence	Probability
[2, 1, 0, 1]	0.0104584
Observation Sequence	Probability
[2, 1, 0, 2]	0.0101768
Observation Sequence	Probability
[2, 1, 1, 0]	0.0118624
Observation Sequence	Probability
[2, 1, 1, 1]	0.0121312
Observation Sequence	Probability
[2, 1, 1, 2]	0.0131264
Observation Sequence	Probability
[2, 1, 2, 0]	0.0120128
Observation Sequence	Probability
[2, 1, 2, 1]	0.0134504
Observation Sequence	Probability
[2, 1, 2, 2]	0.0148168
Observation Sequence	Probability
[2, 2, 0, 0]	0.01546076
Observation Sequence	Probability
[2, 2, 0, 1]	0.01082948
Observation Sequence	Probability
[2, 2, 0, 2]	0.01059376
Observation Sequence	Probability
[2, 2, 1, 0]	0.013118
Observation Sequence	Probability
[2, 2, 1, 1]	0.0136592
Observation Sequence	Probability
[2, 2, 1, 2]	0.0148348
Observation Sequence	Probability
[2, 2, 2, 0]	0.01358644
Observation Sequence	Probability
[2, 2, 2, 1]	0.01536292
Observation Sequence	Probability
[2, 2, 2, 2]	0.01695464

('Final Sum of Probabilities is: ', '1.0')

As we can see, the resultant sum of all possibilities of **observation sequences sums to be 1.**

b) Using the α-pass algorithm, using the code attached (AlphaPass.py) the following result was obtained:

Observation Sequence Alpha Probability
[0, 0, 0, 0] 0.02840236
Observation Sequence Alpha Probability
[0, 0, 0, 1] 0.01835428
Observation Sequence Alpha Probability
[0, 0, 0, 2] 0.01744736
Observation Sequence Alpha Probability
[0, 0, 1, 0] 0.0146524
Observation Sequence Alpha Probability
[0, 0, 1, 1] 0.0131824
Observation Sequence Alpha Probability
[0, 0, 1, 2] 0.0138572
Observation Sequence Alpha Probability
[0, 0, 2, 0] 0.01260644
Observation Sequence Alpha Probability
[0, 0, 2, 1] 0.01300292
Observation Sequence Alpha Probability
[0, 0, 2, 2] 0.01409464
Observation Sequence Alpha Probability
[0, 1, 0, 0] 0.0147616
Observation Sequence Alpha Probability
[0, 1, 0, 1] 0.0099688
Observation Sequence Alpha Probability
[0, 1, 0, 2] 0.0096296
Observation Sequence Alpha Probability
[0, 1, 1, 0] 0.0102496
Observation Sequence Alpha Probability
[0, 1, 1, 1] 0.0101728
Observation Sequence Alpha Probability
[0, 1, 1, 2] 0.0109376
Observation Sequence Alpha Probability
[0, 1, 2, 0] 0.0099968
Observation Sequence Alpha Probability
[0, 1, 2, 1] 0.0110024
Observation Sequence Alpha Probability
[0, 1, 2, 2] 0.0120808
Observation Sequence Alpha Probability
[0, 2, 0, 0] 0.01274924
Observation Sequence Alpha Probability
[0, 2, 0, 1] 0.00880052
Observation Sequence Alpha Probability
[0, 2, 0, 2] 0.00856624

Observation Sequence	Alpha Probability
[0, 2, 1, 0]	0.010022
Observation Sequence	Alpha Probability
[0, 2, 1, 1]	0.0102608
Observation Sequence	Alpha Probability
[0, 2, 1, 2]	0.0111052
Observation Sequence	Alpha Probability
[0, 2, 2, 0]	0.01016356
Observation Sequence	Alpha Probability
[0, 2, 2, 1]	0.01138708
Observation Sequence	Alpha Probability
[0, 2, 2, 2]	0.01254536
Observation Sequence	Alpha Probability
[1, 0, 0, 0]	0.0194936
Observation Sequence	Alpha Probability
[1, 0, 0, 1]	0.0126728
Observation Sequence	Alpha Probability
[1, 0, 0, 2]	0.0120736
Observation Sequence	Alpha Probability
[1, 0, 1, 0]	0.01052
Observation Sequence	Alpha Probability
[1, 0, 1, 1]	0.009632
Observation Sequence	Alpha Probability
[1, 0, 1, 2]	0.010168
Observation Sequence	Alpha Probability
[1, 0, 2, 0]	0.0092584
Observation Sequence	Alpha Probability
[1, 0, 2, 1]	0.0096712
Observation Sequence	Alpha Probability
[1, 0, 2, 2]	0.0105104
Observation Sequence	Alpha Probability
[1, 1, 0, 0]	0.0142016
Observation Sequence	Alpha Probability
[1, 1, 0, 1]	0.0097568
Observation Sequence	Alpha Probability
[1, 1, 0, 2]	0.0094816
Observation Sequence	Alpha Probability
[1, 1, 1, 0]	0.01088
Observation Sequence	Alpha Probability
[1, 1, 1, 1]	0.011072
Observation Sequence	Alpha Probability
[1, 1, 1, 2]	0.011968
Observation Sequence	Alpha Probability
[1, 1, 2, 0]	0.0109504
Observation Sequence	Alpha Probability
[1, 1, 2, 1]	0.0122272

Observation Sequence Alpha Probability
[1, 1, 2, 2] 0.0134624
Observation Sequence Alpha Probability
[1, 2, 0, 0] 0.0140728
Observation Sequence Alpha Probability
[1, 2, 0, 1] 0.0098344
Observation Sequence Alpha Probability
[1, 2, 0, 2] 0.0096128
Observation Sequence Alpha Probability
[1, 2, 1, 0] 0.0118
Observation Sequence Alpha Probability
[1, 2, 1, 1] 0.012256
Observation Sequence Alpha Probability
[1, 2, 1, 2] 0.013304
Observation Sequence Alpha Probability
[1, 2, 2, 0] 0.0121832
Observation Sequence Alpha Probability
[1, 2, 2, 1] 0.0137576
Observation Sequence Alpha Probability
[1, 2, 2, 2] 0.0151792
Observation Sequence Alpha Probability
[2, 0, 0, 0] 0.01893724
Observation Sequence Alpha Probability
[2, 0, 0, 1] 0.01233652
Observation Sequence Alpha Probability
[2, 0, 0, 2] 0.01176224
Observation Sequence Alpha Probability
[2, 0, 1, 0] 0.0103756
Observation Sequence Alpha Probability
[2, 0, 1, 1] 0.0095536
Observation Sequence Alpha Probability
[2, 0, 1, 2] 0.0100988
Observation Sequence Alpha Probability
[2, 0, 2, 0] 0.00919796
Observation Sequence Alpha Probability
[2, 0, 2, 1] 0.00964628
Observation Sequence Alpha Probability
[2, 0, 2, 2] 0.01049176
Observation Sequence Alpha Probability
[2, 1, 0, 0] 0.0151648
Observation Sequence Alpha Probability
[2, 1, 0, 1] 0.0104584
Observation Sequence Alpha Probability
[2, 1, 0, 2] 0.0101768
Observation Sequence Alpha Probability
[2, 1, 1, 0] 0.0118624

Observation Sequence   Alpha Probability
[2, 1, 1, 1]   0.0121312
Observation Sequence   Alpha Probability
[2, 1, 1, 2]   0.0131264
Observation Sequence   Alpha Probability
[2, 1, 2, 0]   0.0120128
Observation Sequence   Alpha Probability
[2, 1, 2, 1]   0.0134504
Observation Sequence   Alpha Probability
[2, 1, 2, 2]   0.0148168
Observation Sequence   Alpha Probability
[2, 2, 0, 0]   0.01546076
Observation Sequence   Alpha Probability
[2, 2, 0, 1]   0.01082948
Observation Sequence   Alpha Probability
[2, 2, 0, 2]   0.01059376
Observation Sequence   Alpha Probability
[2, 2, 1, 0]   0.013118
Observation Sequence   Alpha Probability
[2, 2, 1, 1]   0.0136592
Observation Sequence   Alpha Probability
[2, 2, 1, 2]   0.0148348
Observation Sequence   Alpha Probability
[2, 2, 2, 0]   0.01358644
Observation Sequence   Alpha Probability
[2, 2, 2, 1]   0.01536292
Observation Sequence   Alpha Probability
[2, 2, 2, 2]   0.01695464
('Final Sum of Probabilities is: ', '1.0')

Hence, we can confirm even the sum for all possible probabilities using α-pass is also 1.

Q.4.) a) N=2 hidden states. In this result, we can see that the probabilities of consonants and vowels are easily differentiable. This is because, those are the letters which are the most frequent in an English Text. This was found after using the HMM techniques of passing the data through α-pass, β-pass and gamma's and di-gammas. Finally, according to the algorithm, these values are re-estimated to find a local maximum so that we can find the most likelihood of the hidden states.

Initial A Matrix
0.000059   0.999941
-0.755605 -0.458650
Initial B Matrix
-0.532767 1.384101
-0.218959 0.476597
-0.047045 0.086192
-0.678865 1.336676

-0.679296	0.596052
-0.934693	1.148830
-0.383502	1.373592
-0.519416	1.799658
-0.830965	0.663431
-0.034572	0.448606
-0.053462	1.784246
-0.529700	1.312304
-0.671149	1.368044
-0.007698	1.183115
-0.383416	0.131993
-0.066842	1.147007
-0.417486	1.606570
-0.686773	0.495223
-0.588977	0.792494
-0.930436	1.391904
-0.846167	0.867529
-0.526929	0.431783
-0.091965	0.499213
-0.653919	0.652401
-0.415999	0.302366
-0.701191	0.883483
-0.910321	1.630085

Initial Pi Matrix

0.937548	0.062452

Final A Matrix

0.168665	0.831328
0.688003	0.312015

Final B Matrix

0.123801	0.000000	a
0.000000	0.023101	b
0.001511	0.045829	c
0.005937	0.045822	d
0.201139	0.000000	e
0.000000	0.037320	f
0.012275	0.018206	g
0.000000	0.074056	h
0.123757	0.000000	i
0.013104	0.010940	j
0.001906	0.005623	k
0.000000	0.057278	l
0.000000	0.030631	m
0.000000	0.105564	n
0.114129	0.000000	o
0.004535	0.023990	p
0.000000	0.001389	q
0.000000	0.075956	r
0.007676	0.078121	s

0.029491	0.095450	t
0.042685	0.001044	u
0.000000	0.011697	v
0.000000	0.019994	w
0.000000	0.002815	x
0.017025	0.007257	y
0.000000	0.001133	z
0.301009	0.226802	}

Final Pi Matrix
0.000000 1.000000

b) N=3. In this case, the model tries to distribute the probabilities of these consonants and vowels across 3 states. Hence, we won't get more accurate differentiation than the previous case but will have some notion of that differentiation.

c) When N=4, the model is driving towards differentiation of each letter. Hence, the probabilities will be more and more distributed as we increase the number of states.

d) When N=26, **then the model is able to differentiate between each letter to a specific state.** Hence, we will see accurate distribution of the probabilities.