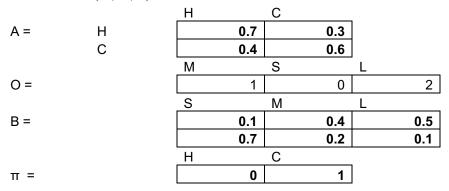
Giang Duong – ID 014533857 Homework I – CS 271 Professor Mark Stamp

1.a We have  $\lambda = (A, B, \pi)$  where



O(1,0,2)

$$P(HHH) = 0*0.4*0.7*0.1*0.7*0.5 = 0$$

$$P(HHC) = 0*0.4*0.7*0.1*0.3*0.1=0$$

$$P(HCH) = 0*0.4*0.7*0.3*0.4*0.5 = 0$$

$$P(HCC) = 0*0.4*0.7*0.3*0.6*0.1 = 0$$

$$P(CHH) = 1*0.2*0.4*0.1*0.7*0.5 = 0.0028$$

$$P(CHC) = 1*0.2*0.4*0.1*0.3*0.1 = 0.00024$$

$$P(CCH) = 1*0.2*0.6*0.7*0.4*0.5 = 0.0168$$

$$P(CCC) = 1*0.2*0.6*0.7*0.6*0.1 = 0.00504$$

The desired probability is the sum of eight probabilities = 0.024880

## b. Compute P using alphal pass

$$\alpha 0(0) = 0.0 * 0.4 = 0$$

$$\alpha 0(1) = 1.0 * 0.2 = 0.2$$

$$\alpha 1(0) = (0 * 0.7 + 0.2 * 0.4) * 0.1 = 0.008$$

$$\alpha 1(1) = (0 * 0.3 + 0.2 * 0.6) * 0.7 = 0.084$$

$$\alpha 2(0) = (0.008 * 0.7 + 0.084*0.4) *0.5 = 0.0196$$

$$\alpha 2(1) = (0.008 * 0.3 + 0.084 * 0.6) * 0.1 = 0.00528$$

$$\Sigma = \alpha 2(0) + \alpha 2(1) =$$
**0.024880**

c. In term of N and T, and couting only mutiplications, what is the work factor for the method in part a? what is the work factor of method b?

For the method a, we need to compute each probabilites each time and its will consume a lot of time and space to stores the caculated value. Requires **2TN**<sup>T</sup>

For the method b, we need to find out the each  $\alpha$  pass. It will be a shortcut to save time and space but they will not have too much details on each probabilities which can be used in future.

#### Requires N2T

2.

a. The best hidden state sequence for dynamic programming sense: CCH

b. The best hidden state in HMM: CCH

Sum the probability of each O(1,0,2):

```
    0
    1
    2

    H
    0
    0.003040
    0.019600

    C
    1
    0.021840
    0.005280
```

3.

a.

```
import math
A = [[0.7, 0.3],
     [0.4, 0.6]]
B = [[0.1, 0.4, 0.5],
     [0.7, 0.2, 0.1]]
pi= [0.6, 0.4]
S = 0; M = 1; L = 2
Set1 = [0, 1, 2]
0 = []
sum = []
totalsum = 0
K = []
for a in range(3):
    for b in range(3):
        for c in range(3):
            for d in range(3):
                 0.append([a, b, c, d])
print(0)
    prob = []
prob.append((pi[0]*B[0][0[i][0]])*(A[0][0]*B[0][0[i][1]])*(A[0][0]*B[0][0[i][2]])*(A[0][0]*B[0][0[i][3]]))
prob.append((pi[0]*B[0][0[i][0]])*(A[0][0]*B[0][0[i][1]])*(A[0][0]*B[0][0[i][2]])*(A[0
][1]*B[1][0[i][3]]))
prob.append((pi[0]*B[0][0[i][0]])*(A[0][0]*B[0][0[i][1]])*(A[0][1]*B[1][0[i][2]])*(A[1
][0]*B[0][0[i][3]]))
    #HHCC
prob.append((pi[0]*B[0][0[i][0]])*(A[0][0]*B[0][0[i][1]])*(A[0][1]*B[1][0[i][2]])*(A[1
][1]*B[1][0[i][3]]))
```

```
prob.append((pi[0]*B[0][0[i][0]])*(A[0][1]*B[1][0[i][1]])*(A[1][0]*B[0][0[i][2]])*(A[0
][0]*B[0][0[i][3]]))
    #HCHC
prob.append((pi[0]*B[0][0[i][0]])*(A[0][1]*B[1][0[i][1]])*(A[1][0]*B[0][0[i][2]])*(A[0
][1]*B[1][0[i][3]]))
prob.append((pi[0]*B[0][0[i][0]])*(A[0][1]*B[1][0[i][1]])*(A[1][1]*B[1][0[i][2]])*(A[1
][0]*B[0][0[i][3]]))
    #HCCC
prob.append((pi[0]*B[0][0[i][0]])*(A[0][1]*B[1][0[i][1]])*(A[1][1]*B[1][0[i][2]])*(A[1
][1]*B[1][0[i][3]]))
    #CHHH
prob.append((pi[1]*B[1][0[i][0]])*(A[1][0]*B[0][0[i][1]])*(A[0][0]*B[0][0[i][2]])*(A[0
][0]*B[0][0[i][3]]))
    #CHHC
prob.append((pi[1]*B[1][0[i][0]])*(A[1][0]*B[0][0[i][1]])*(A[0][0]*B[0][0[i][2]])*(A[0
][1]*B[1][0[i][3]]))
    #CHCH
prob.append((pi[1]*B[1][0[i][0]])*(A[1][0]*B[0][0[i][1]])*(A[0][1]*B[1][0[i][2]])*(A[1
][0]*B[0][0[i][3]]))
    #CHCC
prob.append((pi[1]*B[1][0[i][0]])*(A[1][0]*B[0][0[i][1]])*(A[0][1]*B[1][0[i][2]])*(A[1
][1]*B[1][0[i][3]]))
    #CCHH
prob.append((pi[1]*B[1][0[i][0]])*(A[1][1]*B[1][0[i][1]])*(A[1][0]*B[0][0[i][2]])*(A[0
][0]*B[0][0[i][3]]))
    #CCHC
prob.append((pi[1]*B[1][0[i][0]])*(A[1][1]*B[1][0[i][1]])*(A[1][0]*B[0][0[i][2]])*(A[0
][1]*B[1][0[i][3]]))
    #CCCH
prob.append((pi[1]*B[1][0[i][0]])*(A[1][1]*B[1][0[i][1]])*(A[1][1]*B[1][0[i][2]])*(A[1
][0]*B[0][0[i][3]]))
    #CCCC
prob.append((pi[1]*B[1][0[i][0]])*(A[1][1]*B[1][0[i][1]])*(A[1][1]*B[1][0[i][2]])*(A[1
][1]*B[1][0[i][3]]))
    #print(prob)
    sum = math.fsum(prob)
    totalsum += sum
print("Total Sum = " , totalsum)
```

```
A = [[0.7, 0.3], \\ [0.4, 0.6]]
B = [[0.1, 0.4, 0.5],
[0.7, 0.2, 0.1]]
pi = [0.6, 0.4]
S = 0;
M = 1;
L = 2
Set1 = [0, 1, 2]
[] = 0
totalsum = 0
for a in range(3):
    for b in range(3):
         for c in range(3):
             for d in range(3):
                  0.append([a, b, c, d])
N = len(pi) # total states
M = len(\dot{O}[0]) # total observations
print("M = ",M)
for step in range(81):
    alpha = [[0.0 for i in range(N)] for _ in range(M)]
    for i in range(N):
         alpha[0][i] = (pi[i] * B[i][0[step][0]])
    for t in range(1, M):
         for i in range(N):
             alpha[t][i] = 0
             for j in range(0, N):
                  alpha[t][i] = alpha[t][i] + alpha[t - 1][j] * A[j][i]
             alpha[t][i] = alpha[t][i] * B[i][0[step][t]]
             sum += alpha[M-1][i]
         totalsum += sum
    print("Observation : ", O[step])
print("Alpha = ",alpha)
print("Sum = ", sum)
print("Total Sum = ",totalsum)
```

# **4.** a, N = 2

```
Name: Giang Duong
# Homework 1:
3. Write a HMM program for the English text problem in Section 9.2 of Chapter 9.
Test your each program on each of the following cases.
"""
```

```
import re
import string
import math
import numpy as np
import random
def FABS(x):
 if x < 0.0:
   x = -x
   x = x
def setupfilename():
 # remove all characters except alphabets and word-space, save the new file to new_brown
 with open ("brown.txt", "r") as filename, open ("newbrown.txt", "w") as newline:
   x = filename.read()
   result = re.sub("[^a-z\s]", "", x, 0, re.IGNORECASE | re.MULTILINE)
   result = result.lower()
   newline.write(result)
   filename.close()
 occurrences = {} # count the alphabet
 with open('newbrown.txt', 'r') as file:
   letters = string.ascii_lowercase
   for i in file:
     text_lower = i.lower()
     for letter in letters:
       if letter in text_lower:
          occurrences[letter] = occurrences.get(letter, 0) + 1
   sorted(occurrences)
 # for word in occurrences:
 # stringAZ = list(string.ascii_lowercase +' ')
 count = 0
 with open('newbrown.txt', 'r') as file:
   while (count < countto):</pre>
     c = file.read(1)
     count += 1
       print("End of File")
     char to num = ord(c) - 97
     if (char_to_num < 0):</pre>
        char_to_num = 26 # for white space (DEC = 32); 32-97 < 0 then put it to the last: a-z then space
     stringObservation.append(char_to_num)
 return stringObservation
def getRandom(N):
 r = 1.0 / N
```

```
diff = 0.1 * r
 t = r + (-diff + random.random() * 2 * diff)
 t = np.round(t, decimals=2)
 return t
def rowstocastic(numlist):
 s = sum(numlist)
 norm = [float(i)/s for i in numlist]
 norm = np.round(norm, decimals=2)
 return norm
def createTableA(number_of_state):
 A = \prod
 for i in range(number_of_state):
   A.append([])
   for j in range (number_of_state):
     A[i].append(getRandom(number_of_state))
   A[i] = rowstocastic(A[i])
 print(np.matrix(A))
def createTableB(number_of_state, symbols):
 B = []
 for i in range(number_of_state):
    B.append([])
   for j in range(symbols):
     B[i].append(getRandom(number_of_state))
    B[i] = rowstocastic(B[i])
 print(np.matrix(B))
 return B
def createTablepi(number_of_state):
 for i in range(number_of_state):
   pi.append(getRandom(number_of_state))
 pi = rowstocastic(pi)
 print(pi)
 return pi
Forward algorithm
def forward():
 c[0] = 0.0
 for i in range(number_of_state):
   alpha[0][i] = (pi[i] * B[i][stringObservation[0]])
   c[0] += alpha[0][i]
```

```
c[0] = (1 / c[0])
 for i in range(number_of_state):
   alpha[0][i] = c[0] * alpha[0][i]
 for t in range(1, T):
    c[t] = 0.0
   for i in range(number_of_state):
      alpha[t][i] = 0.0
      for j in range(number_of_state):
        alpha[t][i] = alpha[t][i] + alpha[t - 1][j] * A[j][i]
      alpha[t][i] = alpha[t][i] * B[i][stringObservation[t]]
      c[t] = c[t] + alpha[t][i]
    c[t] = 1 / c[t]
    for i in range(number_of_state):
      alpha[t][i] = c[t] * alpha[t][i]
def backward():
 Tbackward = T - 1
 for i in range(number_of_state):
    beta[Tbackward][i] = 1.0 * c[Tbackward]
 for t in reversed(range(T - 1)):
   for i in range(number_of_state):
      beta[t][i] = 0.0
      for j in range(number_of_state):
        beta[t][i] += (A[i][j] * B[j][stringObservation[t + 1]] * beta[t + 1][j])
      beta[t][i] = c[t] * beta[t][i]
def gammasAnddigammas():
 for t in range(0, T - 1):
    temp2 = 0.0
   for i in range(number_of_state):
      gammas[t][i] = 0
      for j in range(number_of_state):
        digammas[t][i][j] = (alpha[t][i] * A[i][j] * B[j][stringObservation[t+1]] * beta[t+1][j]) # / denom
        gammas[t][i] += digammas[t][i][j]
      temp2 += gammas[t][i]
      temp = 0.0
      for j in range(number_of_state):
        temp += alpha[t][j] * beta[t][j]
      temp = (alpha[t][i] * beta[t][i]) / temp
```

```
if ((FABS(temp - gammas[t][i])) > EPSILON):
       print("gammas ", i, "=", gammas[t][i], temp, "Error!!!")
   if (FABS(1.0 - temp2) > EPSILON):
     print("Sum of gammas's = ", temp2, "should sum to 1.0. \n")
 for i in range(number_of_state):
   gammas[Tbackward][i] = alpha[Tbackward][i]
'"" Re-estimate the model \pi"""
def reestimate():
 for i in range(number_of_state):
   pi[i] = gammas[0][i]
 for i in range(number_of_state):
   for j in range(number_of_state):
     denom = 0.0
     for t in range(Tbackward):
       numer += digammas[t][i][j]
       denom += gammas[t][i]
     A[i][j] = numer / denom
 # Re-estimate matrix B: Note: follow the row stocastic
 for i in range(number_of_state):
   for j in range(symbols):
     numer = 0.0
     denom = 0.0
     for t in range(T):
       if (stringObservation[t] == j):
         numer += gammas[t][i]
       denom += gammas[t][i]
     # if numer == 0:
     B[i][j] = numer / denom
 return A, B, pi
def computeLog(c):
 logProb = 0
 for t in range(T):
   logProb += math.log(c[t])
 return -logProb
```

```
EPSILON = 0.001
symbols = 27 # 26 characters and word-space; M = 27
number of state = 2 # N = 2 because of vowel and consonant
intialA = 1 / number_of_state
intialB = 1 / symbols
Tbackward = T - 1
countfrom = 0
countto = 50000
maxIters = 200
minIters = 20
c = np.zeros((T_i), dtype=np.float32)
alpha = np.zeros((T, number_of_state), dtype=np.float32)
beta = np.zeros((T, number_of_state), dtype=np.float32)
gammas = np.zeros((T, number_of_state), dtype=np.float32)
digammas = np.zeros((T, number_of_state, number_of_state), dtype=np.float32)
stringObservation = []
stringObservation = setupfilename()
A = createTableA(number_of_state)
B = createTableB(number_of_state, symbols)
pi = createTablepi(number_of_state)
iters = 0
logProb = -1.0
newLogProb = 0.0
diff = maxsize
while ((iters < maxIters) and (newLogProb > logProb)):
 print("Iteration: ", iters)
 logProb = newLogProb
  forward()
 backward()
  gammasAnddigammas()
  #print("Gammas, Digammas pass. Done!")
  reestimate()
  logProb = computeLog(c)
  # trick so no initial logProb is requires
  if iters == 0:
   logProb = newLogProb - 1.0
  #print("pi=,", pi)
  print("Iteration = %d, score log [P(observation | lambda)] = %s" % (iters, logProb))
  iters += 1
print("HMM training finished. Model:")
A = np.round(A, decimals=5)
```

```
B = np.round(B, decimals=5)
pi = np.round(pi, decimals=5)
print("Final ==A==", np.matrix(A))
print("Final ==B== ", np.matrix(B))
print("Final ==pi== %s", np.matrix(pi))
```

a. The program genarates final **B^T** matrix. We can see there some B[i][j] in a rows has 0.000 probability which implies there 2 hidden states.

Its implied that there are 2 separate kind of letters in English which is vowels and consonants.

After testing with couple test, the word-space is not has zero probability which means it is not belong to the 2 hidden states(vowels and consonants)

	Final ==B==	
[[	[	
0.14	0.	
0.	0.023	
0.001	0.054	
0.	0.068	
0.215	0.	
0.	0.035	
0.004	0.024	
0.	0.072	
0.123	0.	
0.	0.004	
0.002	0.006	
0.	0.071	
0.	0.038	
0.	0.112	
0.132	0.112	
0.002	0.034	
0.	0.002	
0.	0.1	
0.	0.108	
0.021	0.132	
0.045	0.132	
0.	0.016	
0.	0.022	
0.	0.004	
0.005	0.021	
0.	0.001	
0.31]	0.05	]]
	Final ==pi== 9	%s

[[1.0.]]

## **4.b**

N = 3

The program genarates final  $B^T$  matrix. We can see zero probability occurs once(except B[0]) so we cannot claim that there will be 3 hidden states. Also, there only word-spaces is not has zero which means it is not belong to the 2 hidden states(vowels and consonants). Prove it one more time with N=3 that it does not belong to any states.

### **4.c**

N = 4

The program generates the final **B^T** matrix which has a few zero probability in each row which mean they will not implies about having 4 hidden states. Once again, word-space has no zero probability. Implies it does not belong to any hidden states.

Final ==B==

[0.22 0. 0. 0. 0.337 0. 0. 0.001 0.166 0. 0. 0.0. 0. 0.207 0. 0. 0.001 0. 0. 0.053 0. 0. 0.0.015 0. 0.001]

[0. 0.003 0.034 0.016 0. 0. 0.037 0. 0.006 0. 0.007 0.003 0. 0. 0. 0.033 0. 0. 0.073 0.197 0.002 0. 0. 0.0012 0. 0.576]

 $\begin{bmatrix} 0. & 0.01 & 0.034 & 0.103 & 0. & 0.043 & 0.007 & 0. & 0.001 & 0.006 & 0.116 & 0.047 & 0.233 & 0. & 0.01 \\ 0.002 & 0.148 & 0.116 & 0.036 & 0.022 & 0.027 & 0.009 & 0.01 & 0.017 & 0.002 & 0. & ] \end{bmatrix}$ 

#### **4.d**

N = 26

The program generates the final  $\mathbf{B}^{\Lambda}\mathbf{T}$ , took longer than to execute. Matrix implies that there is not a 26 hidden states.

final  $B^T =$ 

- $\begin{array}{c} v \ 0.20611 \ 0.000000 \ 0.000000 \ 0.0000000 \ 0.0000000 \ 0.000000 \ 0.000000 \ 0.000000 \ 0.000000 \ 0.000000 \ 0.$

#### **REFERENCES**

1. M. Stamp, "A revealing introduction to hidden Markov models."

2.