Problem 1

1a) We are asked to find

|  |  |  |  |
| --- | --- | --- | --- |
|  | |  |  |
|  | a | (0.100)(0.400)=0.040 | (0.120000)\*(0.100000)(0.350000)=0.004200 |
| n | (0.400)(0.300)=0.120 | (0.020000)\*(0.400000)(0.500000)=0.004000 |
| o | (0.200)(0.100)=0.020 | (0.120000)\*(0.100000)(0.500000)=0.006000 |
| t | (0.300)(0.100)=0.030 | (0.040000)\*(0.100000)(0.600000)=0.002400 |

Highest prob = 0.006000

1b) Using the same construction:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | |  |  | 3 |
|  | a | 0.02 | 0.0048 | 0.00048 |
| n | 0.04 | 0.0012 | 0.0012 |
| o | 0.02 | 0.024 | 0.00048 |
| t | 0.12 | 0.0012 | 0.00216 |

Highest prob = 0.002160

1c) Using the same construction:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | |  |  | 3 | 4 |
|  | a | 0.02 | 0.0144 | 0.00021 | 0.000035 |
| n | 0.04 | 0.006 | 0.00018 | 0.00036 |
| o | 0.04 | 0.0036 | 0.0018 | 0.000035 |
| t | 0.09 | 0.0012 | 0.000864 | 0.000054 |

Highest prob = 0.000360

2) Using the Viterbi Algorithm, for the first character, the run time is . For the other characters, the run time is . So the total runtime is .

The brute force method is exponential: .

The Viterbi Algorithm is much faster with large k.

3)

4) The HMM-based approach (by simply replacing characters with words in the above formulation) would do very poorly against Google translate for three major reasons. Firstly, this formulation pays no attention to grammar (i.e. bank could be a verb and noun). Secondly, this formulation requires the two sentences be the same length, which is almost never the case in real life. Lastly, the formulation requires the two sentences to have the same structure; there are cases where one sentence must be inverted to be translated to another.

Appendix:

**1a)**

(0.100)(0.400)=0.040

(0.400)(0.300)=0.120

(0.200)(0.100)=0.020

(0.300)(0.100)=0.030

----------------------------------------

(0.040000)\*(0.100000)(0.050000)=0.000200

(0.120000)\*(0.100000)(0.350000)=0.004200 \*

(0.020000)\*(0.100000)(0.100000)=0.000200

(0.030000)\*(0.100000)(0.400000)=0.001200

(0.040000)\*(0.400000)(0.100000)=0.001600

(0.120000)\*(0.400000)(0.050000)=0.002400

(0.020000)\*(0.400000)(0.500000)=0.004000 \*

(0.030000)\*(0.400000)(0.100000)=0.001200

(0.040000)\*(0.100000)(0.250000)=0.001000

(0.120000)\*(0.100000)(0.500000)=0.006000 \*

(0.020000)\*(0.100000)(0.100000)=0.000200

(0.030000)\*(0.100000)(0.400000)=0.001200

(0.040000)\*(0.100000)(0.600000)=0.002400 \*

(0.120000)\*(0.100000)(0.100000)=0.001200

(0.020000)\*(0.100000)(0.300000)=0.000600

(0.030000)\*(0.100000)(0.100000)=0.000300

----------------------------------------

Highest prob = 0.006000

**1b)**

(0.100)(0.200)=0.020

(0.400)(0.100)=0.040

(0.200)(0.100)=0.020

(0.300)(0.400)=0.120

----------------------------------------

(0.020000)\*(0.100000)(0.050000)=0.000100

(0.040000)\*(0.100000)(0.350000)=0.001400

(0.020000)\*(0.100000)(0.100000)=0.000200

(0.120000)\*(0.100000)(0.400000)=0.004800 \*

(0.020000)\*(0.100000)(0.100000)=0.000200

(0.040000)\*(0.100000)(0.050000)=0.000200

(0.020000)\*(0.100000)(0.500000)=0.001000

(0.120000)\*(0.100000)(0.100000)=0.001200 \*

(0.020000)\*(0.500000)(0.250000)=0.002500

(0.040000)\*(0.500000)(0.500000)=0.010000

(0.020000)\*(0.500000)(0.100000)=0.001000

(0.120000)\*(0.500000)(0.400000)=0.024000 \*

(0.020000)\*(0.100000)(0.600000)=0.001200 \*

(0.040000)\*(0.100000)(0.100000)=0.000400

(0.020000)\*(0.100000)(0.300000)=0.000600

(0.120000)\*(0.100000)(0.100000)=0.001200

----------------------------------------

(0.004800)\*(0.200000)(0.050000)=0.000048

(0.001200)\*(0.200000)(0.350000)=0.000084

(0.024000)\*(0.200000)(0.100000)=0.000480 \*

(0.001200)\*(0.200000)(0.400000)=0.000096

(0.004800)\*(0.100000)(0.100000)=0.000048

(0.001200)\*(0.100000)(0.050000)=0.000006

(0.024000)\*(0.100000)(0.500000)=0.001200 \*

(0.001200)\*(0.100000)(0.100000)=0.000012

(0.004800)\*(0.200000)(0.250000)=0.000240

(0.001200)\*(0.200000)(0.500000)=0.000120

(0.024000)\*(0.200000)(0.100000)=0.000480 \*

(0.001200)\*(0.200000)(0.400000)=0.000096

(0.004800)\*(0.300000)(0.600000)=0.000864

(0.001200)\*(0.300000)(0.100000)=0.000036

(0.024000)\*(0.300000)(0.300000)=0.002160 \*

(0.001200)\*(0.300000)(0.100000)=0.000036

----------------------------------------

Highest prob = 0.002160

**1c)**

(0.100)(0.200)=0.020

(0.400)(0.100)=0.040

(0.200)(0.200)=0.040

(0.300)(0.300)=0.090

----------------------------------------

(0.020000)\*(0.400000)(0.050000)=0.000400

(0.040000)\*(0.400000)(0.350000)=0.005600

(0.040000)\*(0.400000)(0.100000)=0.001600

(0.090000)\*(0.400000)(0.400000)=0.014400 \*

(0.020000)\*(0.300000)(0.100000)=0.000600

(0.040000)\*(0.300000)(0.050000)=0.000600

(0.040000)\*(0.300000)(0.500000)=0.006000 \*

(0.090000)\*(0.300000)(0.100000)=0.002700

(0.020000)\*(0.100000)(0.250000)=0.000500

(0.040000)\*(0.100000)(0.500000)=0.002000

(0.040000)\*(0.100000)(0.100000)=0.000400

(0.090000)\*(0.100000)(0.400000)=0.003600 \*

(0.020000)\*(0.100000)(0.600000)=0.001200 \*

(0.040000)\*(0.100000)(0.100000)=0.000400

(0.040000)\*(0.100000)(0.300000)=0.001200

(0.090000)\*(0.100000)(0.100000)=0.000900

----------------------------------------

(0.014400)\*(0.100000)(0.050000)=0.000072

(0.006000)\*(0.100000)(0.350000)=0.000210 \*

(0.003600)\*(0.100000)(0.100000)=0.000036

(0.001200)\*(0.100000)(0.400000)=0.000048

(0.014400)\*(0.100000)(0.100000)=0.000144

(0.006000)\*(0.100000)(0.050000)=0.000030

(0.003600)\*(0.100000)(0.500000)=0.000180 \*

(0.001200)\*(0.100000)(0.100000)=0.000012

(0.014400)\*(0.500000)(0.250000)=0.001800 \*

(0.006000)\*(0.500000)(0.500000)=0.001500

(0.003600)\*(0.500000)(0.100000)=0.000180

(0.001200)\*(0.500000)(0.400000)=0.000240

(0.014400)\*(0.100000)(0.600000)=0.000864 \*

(0.006000)\*(0.100000)(0.100000)=0.000060

(0.003600)\*(0.100000)(0.300000)=0.000108

(0.001200)\*(0.100000)(0.100000)=0.000012

----------------------------------------

(0.000210)\*(0.100000)(0.050000)=0.000001

(0.000180)\*(0.100000)(0.350000)=0.000006

(0.001800)\*(0.100000)(0.100000)=0.000018

(0.000864)\*(0.100000)(0.400000)=0.000035 \*

(0.000210)\*(0.400000)(0.100000)=0.000008

(0.000180)\*(0.400000)(0.050000)=0.000004

(0.001800)\*(0.400000)(0.500000)=0.000360 \*

(0.000864)\*(0.400000)(0.100000)=0.000035

(0.000210)\*(0.100000)(0.250000)=0.000005

(0.000180)\*(0.100000)(0.500000)=0.000009

(0.001800)\*(0.100000)(0.100000)=0.000018

(0.000864)\*(0.100000)(0.400000)=0.000035 \*

(0.000210)\*(0.100000)(0.600000)=0.000013

(0.000180)\*(0.100000)(0.100000)=0.000002

(0.001800)\*(0.100000)(0.300000)=0.000054 \*

(0.000864)\*(0.100000)(0.100000)=0.000009

----------------------------------------

Highest prob = 0.000360