Assignment - 3

How to compile and run?

- Run using following command:
  python viterbi.py probs.txt sents.txt

Results and Analysis:

For the sentence (mark has fish), the results are as follows:

PROCESSING SENTENCE: mark has fish

FINAL VITERBI NETWORK
P(mark=inf)=1e-08
P(mark=verb)=0.006
P(mark=noun)=0.072
P(mark=prep)=1e-08
P(has=inf)=1.32e-07
P(has=verb)=0.001404
P(has=noun)=4.62e-07
P(has=prep)=2.16e-06
P(fish=inf)=3.0888e-08
P(fish=verb)=2.1021e-08
P(fish=noun)=8.64864e-05
P(fish=prep)=3.51e-08

FINAL BACKPTR NETWORK
BackPtr(fish=inf) = verb
BackPtr(fish=verb) = noun
BackPtr(fish=noun) = verb
BackPtr(fish=prep) = verb

BEST TAG SEQUENCE HAS PROBABILITY = 4.32432e-05
fish -> noun
has -> verb
mark -> noun

FORWARD ALGORITHM RESULTS

P(mark=inf) = 1e-08
P(mark=verb) = 0.006
P(mark=noun) = 0.072
P(mark=prep) = 1e-08
P(has=inf) = 1.327200002e-07
P(has=verb) = 0.00140401822503
P(has=noun) = 4.627208501e-07
P(has=prep) = 2.3100000002e-06
P(fish=inf) = 3.08884300051e-08
P(fish=verb) = 3.78658962653e-08
P(fish=noun) = 8.66446074254e-05
P(fish=prep) = 3.51143616785e-08

For the sentence (mark bears fish), the results are as follows:

FINAL VITERBI NETWORK
P(mark=inf)=1e-08
P(mark=verb)=0.006
P(mark=noun)=0.072
P(mark=prep)=1e-08
P(bears=inf)=1.32e-07
P(bears=verb)=0.000936
P(bears=noun)=9.24e-05
P(bears=prep)=2.16e-06
P(fish=inf)=2.0592e-08
P(fish=verb)=4.2042e-06
P(fish=noun)=5.76576e-05
P(fish=prep)=2.34e-08

FINAL BACKPTR NETWORK
BackPtr(fish=inf) = verb
BackPtr(fish=verb) = noun
BackPtr(fish=noun) = verb
BackPtr(fish=prep) = verb

BEST TAG SEQUENCE HAS PROBABILITY = 2.88288e-05
fish -> noun
bears -> verb
mark -> noun

FORWARD ALGORITHM RESULTS

P(mark=inf) = 1e-08
P(mark=verb) = 0.006
P(mark=noun) = 0.072
P(mark=prep) = 1e-08
P(bears=inf) = 1.327200002e-07
P(bears=verb) = 0.00093601215002
P(bears=noun) = 9.254417002e-05
P(bears=prep) = 2.3100000002e-06
P(fish=inf) = 2.05932171693e-08
P(fish=verb) = 4.22429579097e-06
P(fish=noun) = 5.78161698564e-05
P(fish=prep) = 2.61766532783e-08

For the sentence (mark likes to fish for fish), the results are as follows:

PROCESSING SENTENCE: mark likes to fish for fish

FINAL VITERBI NETWORK
P(mark=inf)=1e-08
P(mark=verb)=0.006
P(mark=noun)=0.072
P(mark=prep)=1e-08
P(likes=inf)=1.32e-07
P(likes=verb)=4.68e-06
P(likes=noun)=4.62e-07
P(likes=prep)=2.16e-06
P(to=inf)=1.019304e-06
P(to=verb)=3.003e-11
P(to=noun)=3.6036e-10
P(to=prep)=3.861e-07
P(fish=inf)=1.019304e-14

P(fish=verb)=5.351346e-08
P(fish=noun)=2.62548e-08
P(fish=prep)=1.08108e-14
P(for=inf)=1.17729612e-12
P(for=verb)=1.706562e-12
P(for=noun)=4.12053642e-12
P(for=prep)=3.07702395e-09
P(fish=inf)=3.7544364e-17
P(fish=verb)=1.8748440711e-13
P(fish=noun)=2.092376286e-10
P(fish=prep)=1.236160926e-16

FINAL BACKPTR NETWORK
BackPtr(fish=inf) = verb
BackPtr(fish=verb) = noun
BackPtr(fish=noun) = prep
BackPtr(fish=prep) = noun

BEST TAG SEQUENCE HAS PROBABILITY = 1.046188143e-10
fish -> noun
for -> prep
fish -> verb
to -> inf
likes -> verb
mark -> noun

FORWARD ALGORITHM RESULTS

P(mark=inf) = 1e-08
P(mark=verb) = 0.006
P(mark=noun) = 0.072
P(mark=prep) = 1e-08
P(likes=inf) = 1.327200002e-07
P(likes=verb) = 4.6800607501e-06
P(likes=noun) = 4.627208501e-07
P(likes=prep) = 2.3100000002e-06
P(to=inf) = 1.01960487002e-06
P(to=verb) = 4.0100755879e-11
P(to=noun) = 5.56720632183e-10
P(to=prep) = 4.31994985803e-07

P(fish=inf) = 1.54037823939e-14
P(fish=verb) = 5.35576107102e-08
P(fish=noun) = 2.93862905339e-08
P(fish=prep) = 3.22201364207e-14
P(for=inf) = 1.17856129901e-12
P(for=verb) = 1.91064561642e-12
P(for=noun) = 4.12423262646e-12
P(for=prep) = 5.10721666377e-09
P(fish=inf) = 9.31593981381e-17
P(fish=verb) = 2.85290943867e-13
P(fish=noun) = 3.47408471329e-10
P(fish=prep) = 2.22577071455e-16

For the sentence (bears fish), the results are as follows:

PROCESSING SENTENCE: bears fish

FINAL VITERBI NETWORK
P(bears=inf)=1e-08
P(bears=verb)=0.002
P(bears=noun)=0.016
P(bears=prep)=1e-08
P(fish=inf)=4.4e-08
P(fish=verb)=0.000728
P(fish=noun)=0.0001232
P(fish=prep)=4.8e-07

FINAL BACKPTR NETWORK
BackPtr(fish=inf) = verb
BackPtr(fish=verb) = noun
BackPtr(fish=noun) = verb
BackPtr(fish=prep) = noun

BEST TAG SEQUENCE HAS PROBABILITY = 0.000364
fish -> verb
bears -> noun

FORWARD ALGORITHM RESULTS

P(bears=inf) = 1e-08
P(bears=verb) = 0.002
P(bears=noun) = 0.016
P(bears=prep) = 1e-08
P(fish=inf) = 4.41600002e-08
P(fish=verb) = 0.00072801452507
P(fish=noun) = 0.00012332868008
P(fish=prep) = 5.300000002e-07