

Problem 1

A

The time complexity of the naive HMM algorithm is

$$O(L^M)$$

where L is the number of states and M is the number of observations. This is because there are L^M possibilities.

B

The time complexity of Viterbi is

$$O(ML^2)$$

where L is the number of states and M is the number of observations. This is because we consider the best k length prefix ending in each observations (so at each length, we need to consider ending in any state, and when we consider ending in a state, we need to consider coming previously from any state).

C

D

```
FILE = sequenceprediction1.txt
MAXSTATESEQ(77550) = 22222
MAXSTATESEQ(7224523677) = 2222221000
MAXSTATESEQ(505767442426747) = 222100003310031
MAXSTATESEQ(72134131645536112267) = 10310310000310333100
MAXSTATESEQ(4733667771450051060253041) = 2221000003222223103222223
P(77550) = 0.00011810020361
P(7224523677) = 2.03268952868e-09
P(505767442426747) = 2.47736684344e-13
P(72134131645536112267) = 8.8711581199e-20
P(4733667771450051060253041) = 3.739956487e-24
```

```
FILE = sequenceprediction2.txt
MAXSTATESEQ(60622) = 11111
MAXSTATESEQ(4687981156) = 2100202111
MAXSTATESEQ(815833657775062) = 021011111111111
MAXSTATESEQ(21310222515963505015) = 0202011111111111021
MAXSTATESEQ(6503199452571274006320025) = 1110202111111102021110211
P(60622) = 2.0883778502e-05
P(4687981156) = 5.18099069122e-11
P(815833657775062) = 3.31515789374e-15
P(21310222515963505015) = 5.12595690734e-20
P(6503199452571274006320025) = 1.29696134676e-25
```

```
FILE = sequenceprediction3.txt
MAXSTATESEQ(13661) = 00021
MAXSTATESEQ(2102213421) = 3131310213
MAXSTATESEQ(166066262165133) = 133333133133100
MAXSTATESEQ(53164662112162634156) = 20000021313131002133
MAXSTATESEQ(1523541005123230226306256) = 1310021333133133313133133
P(13661) = 0.000173176766091
P(2102213421) = 8.28507920611e-09
P(166066262165133) = 1.64163879355e-12
P(53164662112162634156) = 1.06347342033e-16
P(1523541005123230226306256) = 4.53467046372e-22
```

FILE = sequenceprediction4.txt
MAXSTATESEQ(23664) = 01124
MAXSTATESEQ(3630535602) = 0111201112
MAXSTATESEQ(350201162150142) = 011244012441112
MAXSTATESEQ(00214005402015146362) = 11201112412444011112
MAXSTATESEQ(2111266524665143562534450) = 2012012424124011112411124
P(23664) = 0.000114136332244
P(3630535602) = 4.32621950202e-09
P(350201162150142) = 9.79337531929e-14
P(00214005402015146362) = 4.73993465995e-18
P(2111266524665143562534450) = 5.61795840291e-22

FILE = sequenceprediction5.txt
MAXSTATESEQ(68535) = 10111
MAXSTATESEQ(4546566636) = 1111111111
MAXSTATESEQ(638436858181213) = 110111010000011
MAXSTATESEQ(13240338308444514688) = 0001000000011111100
MAXSTATESEQ(0111664434441382533632626) = 211111111111100111110101
P(68535) = 1.32174527463e-05
P(4546566636) = 2.86687793347e-09
P(638436858181213) = 4.32322750273e-14
P(13240338308444514688) = 4.62890402955e-18
P(0111664434441382533632626) = 1.43951249217e-22