 10 Marks (2 + 4 + 4)

1. Consider an HMM with an explicit absorber state w_0 and unique null visible symbol v_0 with the following transition probabilities a_{ij} and symbol probabilities b_{jk} (where the matrix indexes begin at 0):

1	0	0
0.2	0.3	0.5
0.4	0.5	0.1

a_{ij} MATRIX

1	0	0
0	0.7	0.3
0	0.4	0.6

b_{jk} MATRIX

(a) Draw a graphical representation of a hidden Markov model
(b) Suppose the initial hidden state at t = 0 is w_1. Starting from t = 1, what is the probability it generates the particular sequence $V^3 = \{v_2, v_1, v_0\}$?
(a) Given the above sequence V^3, what is the most probable sequence of hidden states?

10 marks (5 + 5)

2. Given:
   a. vocabulary V = { w1,w2,w3}

   b. and the bigram probability distribution p on V X V given by:
      i.   P(w1,w1) = 0.25
      ii. P(w2,w2) = 0.0
      iii. P(w3,w3) = 0.25
      iv. P(w2,w1) = 0.125
      v.   P(w1,w3) =0.25
      vi. P (w1, ?) = 0.5 (i.e. w1 as the first of a pair)
      vii. P (?,w2) = 0.125 (i.e. w2 as the second of a pair)

   Calculate P(w1,w2) and P(w2/w3)

5 marks (2 + 3)

3. We have a language model of vocabulary size 10000. In the training corpus we see donkey 10 times out of which it is followed by "clever" 5 times and "stupid" 5 times.
    a. What is the Maximum Likelihood Estimate of P(clever\donkey)?
    b. What is the Laplace estimate of P(clever\donkey)?