CS281A/Stat241A Homework Assignment 4 (due April 5, 2007)

1. **(IPF)** Consider the undirected graphical model

$$p(x) = \frac{1}{Z} \prod_{(i,j)\in E} \psi_{i,j}(x_i, x_j),$$

with binary variables $x_1, \ldots, x_k$, where the $\psi_{i,j}$ are non-negative functions. The data in the file `hw4-1.data` on the course website consists of n binary vectors of length $k = 5$. Implement the IPF algorithm, and use your implementation on this data to estimate the model parameters for the following graphs:

   (a) $E = \{(1,2), (2,3), (3,4), (4,5), (5,1)\}$,

   (b) $E = \{(1,2), (1,3), (2,4), (3,4), (3,5)\}$.

   Which model fits the data better?

2. **(HMMs with mixtures of Poissons)** Suppose we wish to model traffic in a network using an HMM. Consider an HMM with discrete states q_t (from a set of size m) and non-negative discrete observations y_t, where the conditional distribution of y_t given q_t is a mixture of k Poisson distributions.

   (a) Draw a graphical model for this HMM, representing the observation distributions using an additional latent variable.

   (b) Write the expected complete log likelihood for the model and identify the expectations that you need to compute in the E step of the EM algorithm.

   (c) Outline an algorithm for the E step, based on the standard alpha-beta recursion.

   (d) Write the equations for the M step.

3. **(EM for HMMs)**

   (a) Implement the EM algorithm for HMMs with the observation model of Question 2, where $m = 3$ and $k = 2$.

   (b) Use your implementation with the data on the course web site (in the file `hw4-3.data`) to find maximum likelihood parameter estimates. The data file contains a single sample path of the process; do not attempt to estimate the initial state distribution. For initial values of the parameter estimates, set the Poisson parameter $\lambda_{s,i}$ for state s and mixture component i as

$$\lambda_{1,1} = 1, \qquad \lambda_{2,1} = 50, \qquad \lambda_{3,1} = 200,$$
$$\lambda_{1,2} = 5, \qquad \lambda_{2,2} = 100, \qquad \lambda_{3,2} = 300.$$

   and set all of the other initial distributions to be uniform.

   What are the estimated parameters?

   Evaluate the log likelihood on the training (`hw4-3.data`) and test (`hw4-3.test`) data.

   Explain how you compute the log likelihood.

   (c) Fit a mixture of Poissons with $km = 6$ components to the same data.

   What are the parameter estimates?

   Compare its performance on the training and test data with that of the HMM.