Due Date: November 14, 2005

1. Write matlab code to upgrade your softmax regression to handle unlabeled data.

2. Download HW2.zip and then implement the following functions in Matlab:

   (a) `UG = moralize(DAG)`

   (b) `order = elimOrderGreedy(UG, sizes)`. The algorithm should try to eliminate the first simplicial node (one that is connected to all its uneliminated neighbors, so that no fill-in edges are necessary); if there are no such nodes, it should eliminate the node that results in an induced clique of minimal weight, where the weight of a clique is the product of the sizes of the nodes it contains: $w(C) = \Pi_{i \in C} s(i)$ where $s(i) =$ sizes(i) is the number of values node i can take on.

   (c) `[GT, cliques, fillIns] = triangulate(UG, order)` that triangulates an undirected graph with the specified order. This returns the triangulated version, the maximal cliques, and the fill-in edges.

   (d) `J = jtreeFromMaxCliques(cliques)` that builds a junction tree from the maximal cliques of a chordal graph. You may use the provided function minSpanTree.

3. A common modification of the HMM involves using mixture models for the emission probabilities $p(y_t|q_t)$. For concreteness, let's assume that the y_t are real-valued vectors, and thus our model involves a mixture of Gaussians for each value of the state.

   (a) Draw the graphical model for this modified HMM, identifying clearly the additional latent variables that are needed.

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

   (c) Outline an algorithm for computing the E step, relating it to the standard alpha and beta recursions.

   (d) Write down the equations that implement the M step.