Statistical Methods for Artificial Intelligence, Autumn 2006
Problem set 3, Due Wednesday Oct. 18

Problem 1. Consider a Bayesian network with three variables X_1, X_2, X_3 where X_1 and X_2 have no parents and X_3 has parents X_1 and X_2 so that we have the following equation.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = T_1(x_1)T_2(x_2)T_3(x_3, x_1, x_2) \qquad (1)$$

a) Prove that X_1 and X_2 are independent, i.e., prove the following.

$$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1)P(X_2 = x_2) \qquad (2)$$

Now consider a Markov random field on X_1, X_2, and X_3 with hyperdges $\{X_1\}$, $\{X_2\}$, and $\{X_1, X_2, X_3\}$ so that we have the following.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{1}{Z}e^{-\beta E_1(x_1)}e^{-\beta E_2(x_2)}e^{-\beta E_3(x_1,x_2,x_3)} \qquad (3)$$

Note the similarity between (1) and (3). Suppse that each of the variables X_1, X_2 and X_3 only take on values in the two element set $\{-1, 1\}$.

b) Given an example of the energy functions E_1, E_2 and E_3 for which equation (2) does *not* hold.

Problem 2. Consider an image defined by an $n \times m$ array I of integers in the range 0 to 255. A segmentation of an image is a division of the image into segments. We assume that each segment has an identifier which is also an integer in the range from 0 to 256. Let $I(x, y)$ be the image value at coordinates x, y and let $S(x, y)$ be the segment index for the segment containing pixel x, y. Suppose that we want to find a segmentation S minimizing the following energy where $[\Phi]$ is the indicator function for Φ, i.e., $[\Phi] = 1$ if Φ is true and $[\Phi] = 0$ if Φ is false.

$$E(I, S) = \left(\sum_{x,y}(I(x, y) - S(x, y))^2\right) + \left(\sum_{x,y,\delta\in\{-1,1\},\gamma\in\{-1,1\}} \lambda[S(x, y) \neq S(x + \delta, y + \gamma)]\right)$$

λ is a parameter of the energy function called a "regularization parameter". Explain the effect of increasing or decreasing λ. Also formalize the problem of minimizing this energy as a Markov random field problem. What are the variables, the hyperedges, and the energy functions associated with each hyperedge?

Problem 3. A "series parallel" graph is a graph with special structure (defined below) and with two distinguished nodes called the "interface nodes" of the graph. Series parallel graphs can be defined recursively as follows.

- **edge:** A two node graph with one edge between the two nodes is a seriels parallel graph where the two given nodes are the interface nodes.
- **series combination:** If G_1 is a series parallel graph with interface nodes A and B and G_2 is a series parallel graph with interface nodes B and C, and B is the only node in both G_1 and G_2, then $G_1 \cup G_2$ is a series parallel graph with interface nodes A and C.
- **parallel combination:** If G_1 is a series parallel graph with interface nodes A and B and G_2 is a series parallel graph which also has interface nodes A and B, and A and B are the only node in both G_1 and G_2, then $G_1 \cup G_2$ is a series parallel graph with interface nodes A and B.

a) Draw a series-parallel graph that is a parallel combination of two series combinations of edges.

b) Consider a Markov random field (MRF) where each edge of the graph in part a) is considered to be a hyperedge of the field. Draw a junction tree of width 2 for this MRF.

c) Prove, by structural induction on the definition of a series-parallel graph, that for every series-parallel graph there exists a width 2 junction tree.