UC Berkeley
Department of Electrical Engineering and Computer Science
Department of Statistics

EECS 281A / STAT 241A Statistical Learning Theory

Problem Set 4
Fall 2012

Issued: Tues. Oct. 2, 2012 **Due:** Tues. Oct. 16, 2012

Reading: Chapters 4, 17

Problem 4.1

Sum-product algorithm: Consider the sum-product algorithm on an undirected tree with compatibility functions ψ_s and ψ_{st}. For the tree in Figure 1(a), suppose that we have ternary random variables (i.e., $X_s \in \{-1, 0, 1\}$), where the compatibility functions are of the form:

$$\psi_{st}(x_s, x_t) = \begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}, \quad \psi_s(x_s) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ for } s \text{ odd and,} \quad \psi_s(x_s) = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \text{ for } s \text{ even}$$

(Here entry (i, j) in the 3×3 matrix notation for ψ_{st} gives the value $\psi_{st}(i, j)$, whereas entry i in the 3-vector ψ_s gives the number $\psi_s(i)$).

Throughout this problem, use the serial ordering of message-passing, in which node s updates its message to t only when it has received all other incoming messages.

(a) Implement the sum-product algorithm for a general tree. Please document (meaning describe in comments within the code) what each step is doing, and hand in your documented code. Use it to compute marginal distributions for the tree in Figure 1(a), using the specified compatibility functions with

  (i) $a = 1$ and $b = 0.5$

  (ii) $a = 1$ and $b = 2$.

(b) For any tree, prove that sum-product correctly computes the singleton marginals in at most tree diameter iterations. (The tree diameter is the length of the longest path between any two nodes.)

(c) Show that for each edge $(s,t) \in E$, the message fixed point M^* can be used to compute the pairwise joint distribution over (x_s, x_t) as follows:

$$p(x_s, x_t) \propto \psi_s(x_s)\, \psi_t(x_t)\, \psi_{st}(x_s, x_t) \prod_{u \in N(s) \backslash t} M^*_{us}(x_s) \prod_{u \in N(t) \backslash s} M^*_{ut}(x_t). \tag{1}$$

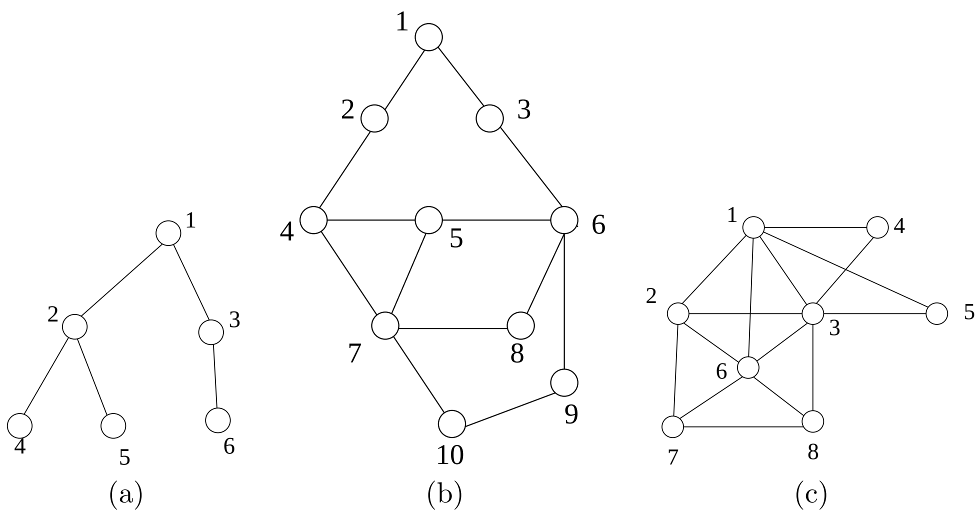

Figure 1: (a) Undirected tree (for Problem 4.1(a)). (b), (c) Some undirected graphs to triangulate (Problem 4.4).

Problem 4.2

Undirected trees and marginals: Let $G = (V, E)$ be an undirected graph. For each vertex $i \in V$, let μ_i be a strictly positive function such that $\sum_{x_i} \mu_i(x_i) = 1$. For each edge, let μ_{ij} be a strictly positive function such that $\sum_{x_i} \mu_{ij}(x_i, x_j) = \mu_j(x_j)$ for all x_j, and $\sum_{x_j} \mu_{ij}(x_i, x_j) = \mu_i(x_i)$ for all x_i. Given integers $k_1, \ldots, k_d$, consider the function

$$\mathbb{Q}(x_1, \ldots, x_d) \quad = \quad \prod_{i=1}^{d} \left[\mu_i(x_i)\right]^{k_i} \prod_{(i,j) \in E} \mu_{ij}(x_i, x_j).$$

Supposing that G is a tree, can you give choices of integers $k_1, \ldots, k_d$ for which $\mathbb{Q}$ is a valid probability distribution? If so, prove the validity. (*Hint:* It may be easiest to first think about a Markov chain.)

Problem 4.3

Max-marginals: For each $i \in \{1, 2, \dots, d\}$, we define the max-marginal at node i via

$$q_i(x_i) \quad = \quad \max_{x_j, j \neq i} \mathbb{P}(x_1, x_2, \dots, x_d).$$

When x_i takes on m-states, then q_i specifies a vector of m numbers. This is the analog of an ordinary marginal distribution, with the summation replaced by maximization.

Suppose that $\arg\max_{x_i} q_i(x_i) = \{x_i^*\}$ for each node (i.e., the maximum is uniquely attained at x_i^*). Show that $(x_1^*, x_2^*, \dots, x_d^*)$ is the unique global optimum, meaning that $\mathbb{P}(x_1^*, \dots, x_d^*) > \mathbb{P}(x_1, \dots, x_d)$ for all $x \neq x^*$.

Problem 4.4

Triangulation/JT: Consider the two graphs shown in panels (b) and (c) of Figure 1. For each graph, first form a triangulated version, and then construct a junction tree using the greedy algorithm. (You can simply implement each step of the greedy algorithm on paper to find a maximum weight spanning tree.)