Homework 3
Due: Oct 14, 2003

1. Consider the function

$$g(x) = ke^{-(x-\theta_1)-(x-\theta_2)^2-(x-\theta_3)^3-\cdots-(x-\theta_d)^d}$$
$$= ke^{-\sum_{i=0}^{d}(x-\theta_i)^i}$$

for some *even* d and $\theta_d \neq 0$.

   (a) Show that there exists a k such that $g(x)$ is a valid PDF. (Hint: you do not need to *find* that k — you merely need to show that it must exist. Consider the behavior of this function for large x, as compared to the behavior for other PDFs that you've seen.)

   (b) Find a maximum likelihood estimator for θ_i. You may leave your answer in the form "the roots of the following polynomial…" without actually finding the roots explicitly, but you must give the correct polynomial.

2. Suppose that you're trying to build a model of the air speed of an unladen swallow (african). You're happy using a Gaussian distribution for your model, and you have strong prior belief that the mean air speed is, itself, Gaussianly distributed with mean 50 KPH. (That is, you believe that the mean airspeed is probably around 50 KPH, but it could be higher or lower, but it's unlikely to be as high as, say, Mach 2.) Your belief is so strong, in fact, that you're willing to assign a standard deviation of no more than 10 KPH to it.

   (a) Let μ and σ be the mean and variance of the swallow's airspeed and ν and τ be the parameters of your prior belief about the swallow's mean airspeed. Give the maximum likelihood (ML) and maxium a posteriori (MAP) estimates for μ, given ν and τ. You may assume that σ is fixed and known.

   (b) Suppose that you take a radar gun (these are metallic african swallows) and measure the airspeed of nine swallows as $X = \{106, 85, 77, 81, 100, 53, 100, 62, 73\}$. What are the respective ML and MAP estimates of μ after you see this data, given the prior beliefs $\nu = 50$ and $\tau = 10$? What about if $\nu = 50$ and $\tau = 1$, $\nu = 50$ and $\tau = 150$, or $\nu = 125$ and $\tau = 5$? In each case, plot (on a single plot) the prior belief, the ML estimate, and the MAP estimate.

   (c) What is the effect of the prior as $N \Rightarrow \infty$ (i.e., as the amount of *observed* data grows)?

   (d) Perhaps a Gaussian prior belief is too strong. It does, for example, allow a negative value for μ or (with small probability) of Mach 2. Maybe we should prefer a uniform prior. Now assume a uniform prior over the range $[\tau_l, \tau_h]$ and derive the form of the MAP estimate for μ.

3. For each of the following Bayesian networks,

   (i) Give the factored form of the joint distribution and the number of parameters in the CPTs (assume that each variable is discrete and can take on k values; an asymptotic/$O()$ result is acceptable).

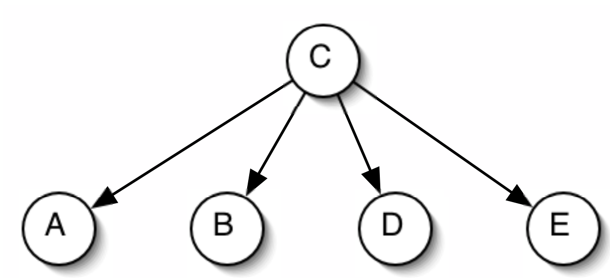
(ii) Give the moralized graph.

(iii) Give a triangulated graph.

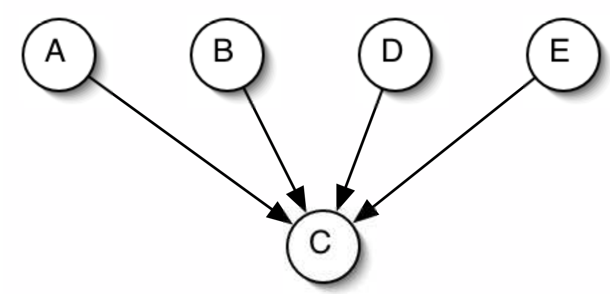
(iv) Give the junction tree.

(v) Describe the sequence of eliminations/messages (i.e., marginalizations and joint probability tables) necessary to calculate $p(C)$ given evidence at A and E.

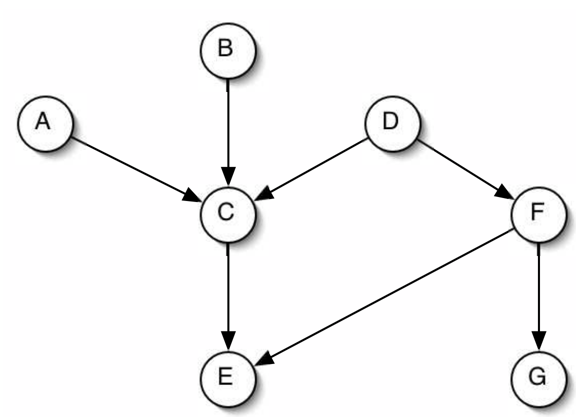
(a)



(b)



(c)



(d)