Homework in Bayesian Modeling and Bayesian Networks

1. Consider the hypotheses S and G in the hypothesis space of "rectangles" in Fig. 2.4 of the textbook (Alpaydin, p.22), S is the most specific hypothesis in the space and G is the most general hypothesis in the space. That is, S gives the "tightest" rectangle that includes all the positive examples and none of the negative examples. The actual class C may be larger than S but is never smaller. On the other hand, G gives the largest rectangle we can draw that includes all the positive examples and none of the negative examples. The actual class C can be smaller than G but is never larger. Clearly, a case of noise-free data is assumed.
   Please use the idea of risk minimization, with the zero-one loss (3.7, 3.9 in the textbook, p.43) to explain why the average of S and G can be a good choice for the final classifier? The average of S and G is the one in the middle of S and G. If S denotes the classifier $p_1 \leq x_1 \leq p_2$ combined with $e_1 \leq x_2 \leq e_2$ and G denotes the classifier $q_1 \leq x_1 \leq q_2$ combined with $f_1 \leq x_2 \leq f_2$, then the middle one of them is given by $(p_1 + q_1)/2 \leq x_1 \leq (p_2 + q_2)/2$ combined with $(e_1 + f_1)/2 \leq x_2 \leq (e_2 + f_2)/2$.
   (hint: you can adopt the approach of Bayes optimal classifier, with a uniform prior!)

2. Consider Bayesian networks of four random variables X_1, X_2, X_3 and X_4, please list *all* of the correspondent networks with the given (conditional) independence rules:

   (a) $(X_2, X_3, X_4) \perp X_1$,
   (b) $X_2 \perp (X_3, X_4)$, $X_1 \perp X_4 | X_2, X_3$,
   (c) $X_3 \perp X_4 | X_2$, $X_2 \perp X_1 | X_3, X_4$,
   (d) $X_1 \perp X_2$.

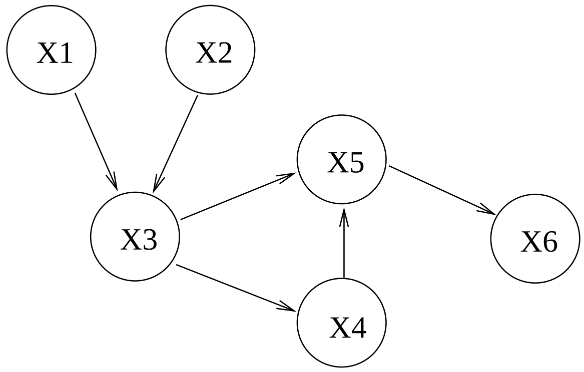

Figure 1: Bayesian network in question 3

3. Given the following Bayesian network with binary random variables, i.e., each random variable has only two states 0 or 1, please compute the marginal probability $P(X_4)$, with no observed variables. Your computation should be *as efficient as possible*. Right now, suppose an evidence is given as $X_5 = 0$, what is your answer again?