

# CSE 250A Quiz 3

Tuesday October 23, 2012

*Instructions.* You should do this quiz in partnership with exactly one other student. Write both your names at the top of this page. Discuss the answer to the question with each other, and then write your joint answer below the question. It is ok if you overhear other students' discussions, because you still need to decide if they are right or wrong. You have seven minutes.

(a) Let  $X_1$  to  $X_d$  and  $Y$  be binary random variables in a Bayesian network where  $Y$  depends on  $X_1$  to  $X_d$  via logistic regression. Explain how to calculate  $p(X_1|Y, X_2, \dots, X_d)$  efficiently. What is the big-O time complexity of this calculation?

*Reminders.* The logistic regression model is  $p(Y = 1|X_1 = x_1, \dots, X_d = x_d) = \sigma(\sum_{i=1}^d w_i x_i)$  where the  $\sigma$  function needs only  $O(1)$  time to evaluate. Remember the slogan "the denominator is the sum of the numerators."

*Answer.* Write  $Z$  for  $X_2, \dots, X_d$  and use Bayes' rule:

$$p(X_1|Y, Z) = \frac{p(Y|X_1, Z)p(X_1|Z)}{p(Y, Z)} = \frac{p(Y|X_1, Z)p(X_1)}{p(Y, Z)}.$$

The second factor in the numerator simplifies to  $p(X_1)$  assuming that the  $X_i$  are unconditionally independent, and  $p(X_1)$  can be looked up in the definition of the Bayesian network. The first factor can be evaluated in  $O(d)$  time for each value of  $X_1$ , because it simply uses the logistic regression formula. The denominator can be evaluated as the sum of the numerators for  $X_1 = 0$  and  $X_1 = 1$ . In total, the time complexity is just  $O(d)$ .

*Additional notes.* The question does not say explicitly that the  $X_i$  are unconditionally independent, but it is reasonable to assume that the edges mentioned from each  $X_i$  to  $Y$  are the only edges. The time complexity is  $O(d)$  to compute  $p(X_1|Y, X_2, \dots, X_d)$  for given values of  $X_2$  to  $X_d$ . For all combinations of those values, the time complexity is  $O(d2^{d-1})$  unless further simplification is possible.