STUDENT NAME: _____

STUDENT ID: _____

FIRST EXAMINATION

Probabilistic reasoning in AI - Winter 2006

February 15, 2006

You are allowed one double sided cheat sheet.

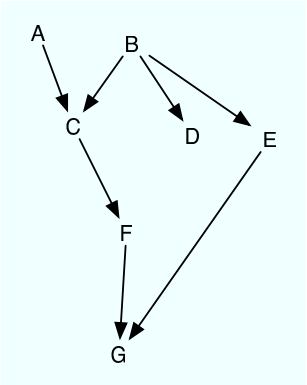
There are 9 questions, for a total of 100 points. Please read the questions first, as they cover a whole range of topics.

Answer all questions on the exam booklet

Good luck!

1. [25 points] **Independence, exact inference**

   Consider the following Bayes network:



   (a) [5 points] For each of the following assertions state whether they are true or false:
      i. $A \perp\!\!\!\perp B$
      ii. $A \perp\!\!\!\perp B|C$
      iii. $A \perp\!\!\!\perp B|D$
      iv. $C \perp\!\!\!\perp E|B$
      v. $A \perp\!\!\!\perp F|C$

   (b) [5 points] Write the parameters on the network picture above. Assuming a tabular representation of the CPDs, and binary variables, how many parameters will there be?

   (c) [2 points] Draw the moral graph of the network

   (d) [3 points] What is the Markov blanket of node C?

(e) [5 points] Draw the junction tree for the the network

(f) [5 points] Write the initial parameters of the junction tree

2. [10 points] **Short questions**

   (a) [5 points] True or false: $p(A) \geq p(A, B), \forall A, B$

   (b) [5 points] You have a Bayes net with 100 nodes. Imagine the lexicographic ordering of the nodes, such that a node comes after all its parents in the ordering. You have a query conditioned on a variable that is almost at the end of this ordering. You want to use approximate inference to compute conditional probability queries for this variable. Should you use likelihood weighting or Gibbs sampling? Justify your answer in one sentence

3. [10 points] **Maximum likelihood estimation**

   Let X be a random variable drawn uniformly from and interval $[a, b]$, where a and b are unknown. More precisely,

$$p(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{if } x > b \end{cases}$$

   You observe a sequence of samples $x_1, \ldots x_n$, with $n \geq 2$. Assume that at least two of these samples have different values.

   Compute the maximum likelihood values for a and b.

4. [5 points] **Undirected models**

   Let X_i, $i = 1 \ldots n$ and Y be random variables. Draw a Markov network such that, for all $i \neq j$, $X_i \perp\!\!\!\perp X_j | Y$.

5. [5 points] **Undirected models** Suppose that someone asks you to draw a Markov network over variables A, B, C and D which obeys the following conditional independency properties: $A \perp\!\!\!\perp B|C$ and $A \perp\!\!\!\perp C|D$. Is this possible? If so, draw the network. If not, justify why.

6. [15 points] **Parameter estimation**

   You are given the following table of data, corresponding to the model:

   $$A \rightarrow C \leftarrow B$$

   Each row corresponds to a patient.

A	B	C
1	1	1
0	0	1
0	0	0
0	0	0
1	1	1
1	0	1
1	1	0
1	1	0
0	1	0
0	0	0

   (a) [5 points] What are the maximum likelihood parameters based on this data?

   (b) [5 points] Is there any problem with doing maximum likelihood estimation here? Would Bayesian learning be better? Justify your answer

   (c) [5 points] Suppose you get a query with $C = 0$. What settings of A and B are possible, and with what probability?

7. [10 points] **Gibbs sampling**

   Suppose you do Gibbs sampling in a Bayes net with no evidence. What does the Markov chain look like? (Give a qualitative description). What will be its stationary distribution?

8. [10 points] **Inference in Markov nets**

   Let X be a node in a Markov network. and let y be an assignment of values to the nodes $Y = MB(X)$. Show that the ratio:

   $$\frac{p(X = x'|y)}{p(X = x|y)}$$

   can be computed efficiently, based only on local parameters. Explain how to find the most likely value of a variable given evidence about the parents based on this observation.

9. [10 points] **Using Bayes nets**

   A car insurance company wants to build a probabilistic model of customer behavior. In the future, they plan to use this in order to determine premiums for new customers, based on the estimated probability of accidents. They hire you as a consultant and present you with a data set of 50 customers, for which they recorded: age (5 brackets), gender (2 values), occupation (5 categories), car make (100 values), number of accidents within a one-year period (integer). They want you to build a model. What do you do?

   (a) Tell them to go get a serious data set and then come find you. There is too little information to do anything.

   (b) Convince a company expert to sit down with you and make up a reasonable model based on their experience. There is too little data to use it.

   (c) Convince a company expert to sit down with you and draw a model structure, then you use maximum likelihood to fit parameters.

   (d) Convince a company expert to sit down with you and draw a model structure, then use Bayesian learning to fit parameters.

   (e) Use the data to do structure learning with simulated annealing and MDL scoring

   (f) None of the above. In this case, state your choice.

   Justify your choice in maximum 3 sentences.