10-601 Machine Learning: Assignment 5

- The assignment is due at 3:00pm (beginning of class) on **Monday, March 3, 2008**.
- Since you only have one week, it will be worth **40 points**.
- Write your name at the top right-hand corner of each page submitted.
- Each student must hand in a hard-copy writeup. See the course webpage for collaboration policies.

Q1: Representing and learning Bayes nets [25 pts]

1. Download JavaBayes from here: `http://www.cs.cmu.edu/~javabayes/`. To run JavaBayes, go into the Classes directory and run `java JavaBayes`.
2. Load the JohnMaryCall network into JavaBayes.
3. Using the "Edit Function" button, look at the CPT's of the various nodes (by clicking on the nodes and then hitting "Dismiss" to close without changing the CPT). Based on these probability tables, what is the joint probability of {**Burglary=True**, **Earthquake=False**, **Alarm=True**, **JohnCall=True**, **MaryCall=True**}?

   We can write the joint probability distribution as:

   $P(Burglary, Earthquake, Alarm, JohnCall, MaryCall) =$
   $P(Burglary)P(Earthquake)P(Alarm|Burglary, Earthquake)$
   $P(JohnCalls|Alarm)P(MaryCalls|Alarm)$

   $P(Burglary = True) = 0.001$
   $P(Earthquake = False) = 0.998$
   $P(Alarm = True|Burglary = True, Earthquake = False) = 0.95$
   $P(JohnCalls = True|Alarm = True) = 0.9$
   $P(MaryCalls = True|Alarm = True) = 0.7$
   Total product is approximately 0.0006.
4. Try querying each node, and note the initial posterior probabilities on each. Now observe **Alarm=True** (by using the Observe button and clickin on **Alarm**). How do the posterior probabilities of the other variables change? Explain why (or why not).

   P(Earthquake= True) becomes approx. 0.77, when it was a very low probability before. Similar change occurs with Burglary. Probabilty of JohnCall and MaryCall also increase. This is because all other nodes are dependent on Alarm.
5. Now, with **Alarm=True**, observe **Earthquake=True**. What happens to the posterior probability of **Burglary**? What happens to the posterior probability of **JohnCall**? Explain the results.

The probability of Burglary decreases (from 0.62 to 0.5) once Earthquake is observed as True. This is an occurrence of "explaining away". Note that Burglary and Earthquake are not conditionally independent, given Alarm.

The probability of JohnCall does not change, as JohnCalls is conditionally independent of Earthquake, given Alarm. Since we already know the Alarm is going off, that is the only variable important in determining JohnCall.

6. Set **Earthquake** to unobserved. Now observe **JohnCall=True**. What happens to the posterior probability of **MaryCall**? What about the posterior probability of **Earthquake**? Explain.

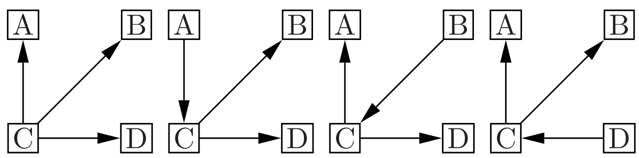
   Nothing in either case. As we mentioned before, JohnCall and Earthquake are conditionally independent given Alarm. Also, JohnCall and MaryCall are conditionally independent given Alarm.

Q2: D-separation [15 pts]

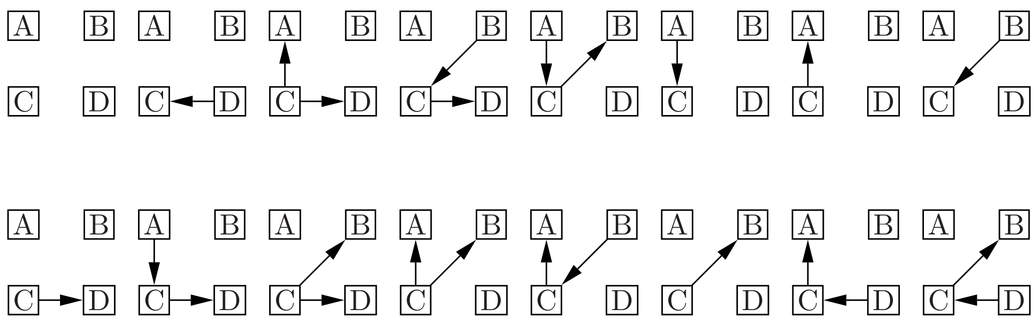
1.

   (a) Draw **all** 4-node Bayesian networks that are consistent with **at least** both of the following conditional independencies (solutions which both allowed or disallowed disconnected components were acceptable): $(A \perp B | C)$, $(D \perp A, B | C)$.

   Connected graphs:



   Disconnected graphs:



   (b) How many different graphs did you draw?

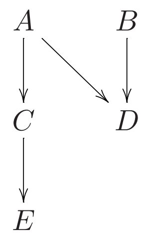
   4 (or 20, allowing for disconnected components).

   (c) Why might you prefer one of these graphs to the others?

   Depending on what the variables are, some configurations might be more intuitive (causality assumptions). The first graph seems simplest, since it's only two levels (a la naive Bayes). The disconnected graphs have various appeals, depending on domain knowledge, intepretation of the variables.

2.

(a) Write all the conditional independencies you can read off this graph in the form $(X \perp Y | \mathbf{Z})$, where X and Y are single variables and $\mathbf{Z}$ is any set of variables, including the empty set:



There are several possibilities for Z in several cases– I have listed optional variables in parentheses where the extra variables will not affect the indepdnencies.

$A \perp B | \emptyset (optionally C, E)$
$A \perp E | C$ (and optionally B, D)
$B \perp C | \emptyset$ (and optionally A, E)
$B \perp E | \emptyset$ (and optionally A, C)
$C \perp D | A$ (and optionally B, E)
$D \perp E | A$ (or C, optionally B)