CS 289 Knowledge representation and reasoning
Spring 1999 Stuart Russell **Assignment 3**

Probabilistic reasoning, due 5/10

This is the final assignment for CS 289 (i.e., it includes the term project). If you have already started on another term project you may substitute it for questions 7 and 8. This assignment may be done in pairs or threes. The deadline of May 10 is a hard deadline.

1. Write an algorithm that computes the truth of an arbitrary first-order logic sentence in a given finite model with a given interpretation for the symbols. (You will need to define a data structure for models and a concrete syntax for FOL.) Show the algorithm working for some nontrivial examples. Find a way to describe its complexity.

2. Let X and Y be two continuous variables; $P(X)$ is a Gaussian distribution with mean μ_X and variance σ^2_X, and $P(Y|X)$ is a linear Gaussian distribution with mean $aX+b$ and variance $\sigma^2_{Y|X}$.
   (a) Show that $P(Y)$ is Gaussian, and compute its mean and variance.
   (b) Show that $P(X|Y)$ is Gaussian, and compute its mean and variance.

3. Let $p_1(x)$ and $p_2(x)$ be two probability density functions, and let $f(x)$ be any monotonically nondecreasing function. Show that if p_1 first-order stochastically dominates p_2, then $E_{p_1}[f(x)] \geq E_{p_1}[f(x)]$.

4. Let X be a node in a belief network and let $MB(X)$ be its Markov blanket. Show that $P(X=x|MB(X))$ is given by
$$P(x|Parents(X)) \times \prod_{Z_j \in Children(X)} P(z_j|Parents(Z_j)).$$
   Does it make sense to precompute or cache these values when using MCMC on a belief network?

5. Show that the Metropolis-Hastings sampler gives detailed balance with the stationary distribution $\pi(x)$, for any proposal distribution q.

6. Show that Gibbs sampling is a special case of Metropolis-Hastings.

7. Implement MCMC for belief net inference. Use the belief net code in
`http://www.cs.berkeley.edu/~russell/code-new/uncertainty/`
and define an interface analogous to those for the other inference methods. Test your code on the oil network in
`http://www.cs.berkeley.edu/~russell/code-new/uncertainty/domains/oil.bn`
by comparing the estimated probability to the true probability computed by, say, enumeration, as a function of the number of iterations.

8. The paper "Probabilistic Frame-Based Systems" by Koller and Pfeffer defines a language for specifying probabilistic knowledge bases that include objects and relations. Consider a variant of the language, in which the entire set of domain objects and the relationships between them are fully specified by the knowledge base, using the closed-world assumption. In other words, for each complex attribute A of a class C, and for each instance X of C, the knowledge base specifies the exact set of instances Y such that $X.A = Y$. In this variant, all complex attributes are multi-valued, but there is no uncertainty as to the number of values. In addition there is no reference uncertainty.

(a) Show that for this variant, if the dependency graph is acyclic, then the entire knowledge base corresponds to a finite belief network.

(b) Write an algorithm for constructing the propositional belief network corresponding to a given knowledge base, given these restrictions. Show it working on the knowledge base given in `http://www.cs.berkeley.edu/~russell/classes/cs289/s99/a3-kb.lisp`

(c) Write an MCMC query-answering system for simple attribute-value queries about named instances, given a knowledge base in the restricted syntax of part (a).

(d) Using your algorithm from Question 1, write an algorithm for approximating the probability of an arbitrary first-order sentence, given a knowledge base.

(e) (Extra credit.) Extend this to handle free variables in the query, as appropriate.