Note:

1. Due 11:59PM, October 10, 2016.
2. Electronic submission to your instructor's email.
3. You are VERY MUCH encouraged to form teams to discuss proofs and program algorithms. If so, please acknowledge your teammate(s)' contributions at the beginning of your submitted homework. You must independently write your homework based on your own understanding.
4. Choose any programming language you like, R, Python, Matlab, C/C++, Julia, etc.

Theory

[10+5 points] Exercise 2.5, Koller and Friedman (2009).
Remark: simple useful fact.

[10+5 points] Exercise 2.9, Koller and Friedman (2009).
Remark: Familiarize yourself with semi-graphoid axioms, especially their intuitions in the setting of probabilistic reasoning. Please refer to Section 1.1.5, Pearl (2009) *Causality*.
Remark 2: Exercise 2.9 differs across versions. For some, it has four sub-problems asking you to prove or disprove conditional independence implications. For others, it asks you to use breadth-first search to determine whether a graph is cyclic. Completing either will count towards the final credit. The second question is more time-consuming but interesting. If you are doing the first question, in 2.9(b), (X,Y ⊥ W | Y) should be (X,Y ⊥ W | ***Z***).

[15+5 points] Exercise 3.13, Koller and Friedman (2009).
Remark: Convince yourself that Theorem 3.5 in the same book is true.

[Now Extra Credit: 15 points] Exercise 7.7, Koller and Friedman (2009).
Remark: In Gaussian setting, it derives a formula for pairwise conditional independence, based on which you can draw an undirected graph. Please give a specific numerical example to illustrate this formula.
Remark 2: If you see a typo: $\mathrm{Cov}_{\mathrm{p}}[X_i; X_j \mid Y]$, change the conditioning variable *Y* to *Z*, which means all random variables except X_i and X_j

[*Extra credit:* 15 points] Prove **Theorem (Hammersley-Clifford-Besag; Discrete Version)**. Suppose that *G = (V,E)* is an undirected graph, where $X_i \in V, j = 1, \dots, d$, are random variables, each taking on a finite number of values. If $P(X_1 = x_1, \dots, X_d = x_d)$ is strictly positive for arbitrary realized values $(x_1, \dots, x_d)'$ and *P* satisfies the local Markov property with respect to *G*, then it factorizes with respect to *G*.
Remark: "fundamental theorem of random fields".

Examples and Implementations

[25 points] Question 3, Professor Eric Xing's course, as linked below:
http://www.cs.cmu.edu/~epxing/Class/10708/hw/pgm_2014_homework1.pdf

[25 points; Choose One: a or b]

a) Question 4 in the linked pdf above. Access data at:
http://www.cs.cmu.edu/~epxing/Class/10708/hw/hw1_images.mat

b) Exercise 3.28, Koller and Friedman (2009). Run your algorithm on your favorite DAG to demonstrate its correctness.

[Extra credit: 15 points] Exercise 3.11, Koller and Friedman (2009)