1. A DAG $\mathcal{D}$ is said to be *perfect* if all parents are married, i.e. if it holds that

$$\alpha, \beta \in \mathrm{pa}(\gamma) \Rightarrow \alpha \rightarrow \beta \text{ or } \beta \rightarrow \alpha.$$

   (a) Show that a perfect DAG $\mathcal{D}$ is Markov equivalent to its *skeleton* i.e. the undirected graph obtained by ignoring directions on all arrows;

   (b) Show the converse, i.e. that if $\mathcal{D}$ is Markov equivalent to its skeleton, the $\mathcal{D}$ is perfect.

   (c) Show that the skeleton of a perfect $\mathcal{D}$ is chordal.

   (d) Show that the edges of a chordal graph $\mathcal{G}$ can be directed to create a Markov equivalent DAG.

   Hint: Exploit the existence of a perfect numbering of a chordal graph.

2. The *conditional entropy* $H(X \mid Y)$ is defined as the average entropy in the conditional distribution

$$H(X \mid Y) = \mathbf{E}[\mathbf{E}\{-\log f(X \mid Y) \mid Y\}] = \sum_y \left\{ \sum_x -f(x \mid y) \log f(x \mid y) \right\} f(y).$$

   (a) Use the information inequality to show that

$$H(X \mid Y) \leq H(X),$$

   i.e. the *entropy is always reduced by conditoning*

   (b) Show that

$$H(X, Y) = H(X \mid Y) + H(Y).$$

   (c) For three discrete random variables, show that

$$H(X, Y, Z) + H(Z) \leq H(X, Z) + H(Y, Z).$$

   (d) Show further that

$$X \perp\!\!\!\perp Y \mid Z \iff H(X, Y, Z) + H(Z) = H(X, Z) + H(Y, Z).$$

3. Consider the DAG $\mathcal{D}$ with arrows $A \rightarrow C, B \rightarrow C, B \rightarrow D, C \rightarrow E, D \rightarrow F, E \rightarrow G, E \rightarrow H, F \rightarrow G, G \rightarrow J, I \rightarrow J$.

   (a) Find the moral graph $\mathcal{D}^m$ of $\mathcal{D}$;

   (b) Find a *minimal chordal cover* $\mathcal{G}$ of $\mathcal{D}^m$, i.e. a chordal graph $\mathcal{G} \supset \mathcal{D}^m$ with the property that removal of any edge in $\mathcal{G}$ which is not an edge in $\mathcal{D}^m$ will not be chordal;

   (c) Arrange the cliques of $\mathcal{G}$ in a junction tree;

   (d) For a specification of all conditional distributions $p_{v \mid \mathrm{pa}(v)}, v \in V$, allocate appropriate potentials to the junction tree to prepare for probability propagation.

4. Consider random variables $X_1, \ldots, X_6$ taking values in $\{-1, 1\}$ and having distribution P with joint probability mass function determined as

$$p(x) \propto \exp\{\theta(x_1x_2 + x_2x_3 + x_3x_4 + x_3x_5)\},$$

where $\theta \neq 0$.

(a) Find the dependence graph of P and identify its cliques;

(b) Set up an appropriate junction tree for probability propagation;

(c) Allocate potentials to cliques;

(d) Calculate $P(X_5 = 1 \mid X_1 = 1, X_4 = 1)$ by probability propagation.