14.7 The Markov blanket of a variable is defined on page 517. Prove that a variable is independent of all other variables in the network, given its Markov blanket and derive Equation (14.12) (page 538).

"A node is conditionally independent of all other nodes in the network, given its parents, children, and children's parents—that is, given its Markov blanket" [AIMA 3E, p.517]

"The probability of a variable given its Markov blanket is proportional to the probability of the variable given its parents times the probability of each child given its respective parents." [AIMA 3E, p.538]

So, we will try to arrive to equation 14.12 that is:

$$P(x'_i \mid mb(X_i)) = \alpha \times P(x'_i \mid parents(X_i)) \times \prod_{Y_j \in Chlldren(X_i)} P(Y_j \mid parents(Y_j))$$

Using full distribution probability with conditional rule:

$$P(x_i \mid x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) = \frac{P(x_1, x_2, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n)}{P(x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n)}$$

Using marginalization rule (on right hand's denominator):

$$P(x_i \mid x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) = \frac{P(x_1, x_2, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n)}{\sum_{x \in X_i} P(x_1, x_2, ..., x_{i-1}, X_i = x, x_{i+1}, ..., x_n)}$$

Using chain rule in a Bayesian Network (on right hand):

$$P(x_i \mid x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) = \frac{\prod_{j=1}^{n} P(x_j \mid parents(X_j))}{\sum_{x \in X_i} \prod_{j=1}^{n} P(x_j \mid parents(X_j))}$$

The terms of the form $P(x_j \mid parents(X_j))$ can be factorized in the numerator and denominator for each term where $x_j \neq x_i \wedge x_j \notin children(X_i)$. I will name all of these terms as β . Then:

$$P(x_i \mid x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) =$$

$$\frac{\beta \times P(x_i \mid parents(X_i)) \times \prod_{y_j \in children(X_i)} P(y_j \mid parents(Y_j))}{\beta \times \sum_{x_i \in X_i} P(x_i \mid parents(X_i)) \times \prod_{y_j \in children(X_i)} P(y_j \mid parents(Y_j))}$$

Then we can suppress β :

$$P(x_i \mid x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) = \frac{P(x_i \mid parents(X_i)) \times \prod_{y_j \in children(X_i)} P(y_j \mid parents(Y_j))}{\sum_{x_i \in X_i} P(x_i \mid parents(X_i)) \times \prod_{y_j \in children(X_i)} P(y_j \mid parents(Y_j))}$$

Where for any $x_i \in X_i$ in the left hand we found that the right hand's denominator value will always be the same. Then we can write the denominator as a constant named: α^{-1} .

$$P(x_i \mid x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) = \alpha \times P(x_i \mid parents(X_i)) \times \prod_{y_j \in \, children(X_i)} P(y_j \mid parents(Y_j))$$

So, looking at the last formula we can see that the only required node variables to resolve the equation are the ones that belongs to $parents(X_i)$, $children(X_i)$ or $parents(children(X_i))$. And these group of variables conform what we call the Markov Blanket of node X_i . Then we can rewrite the last equation as:

$$P(x_i \mid mb(x_i)) = \alpha \times P(x_i \mid parents(X_i)) \times \prod_{y_j \in \, children(X_i)} P(y_j \mid parents(Y_j))$$

Q.E.D.