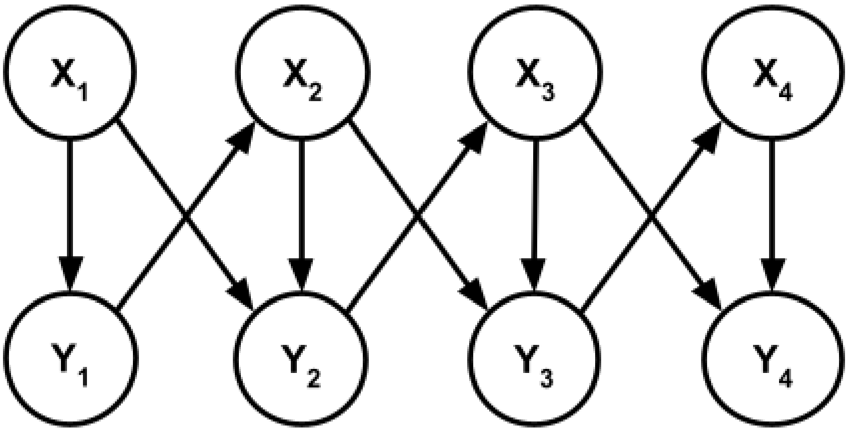
Question 1 - Bayesian Networks (30 points, 20 minutes)

Consider the following Bayesian Network on variables $\mathbf{X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3, Y_4}$.



A) Write the joint distribution in terms of the CPDs. (6 points)

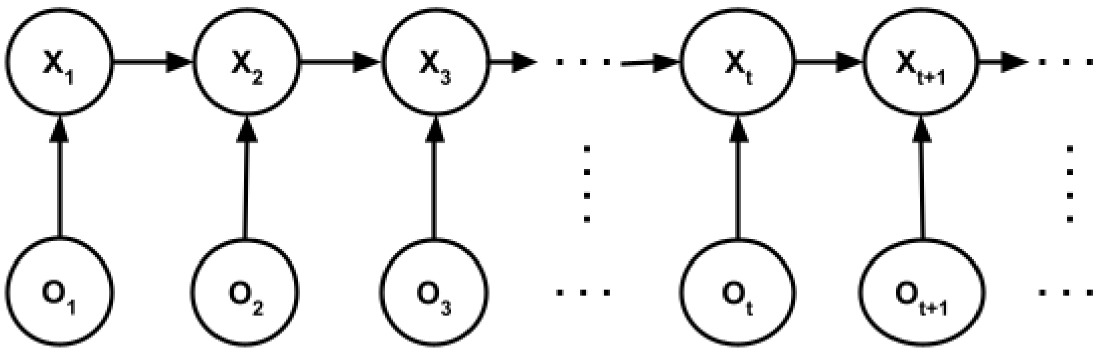
$\mathbf{P(X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3, Y_4) =}$

B) Which of the following statements are **True**, and which are **False** (in general). For each statement, write the word "True" or an Active Trail rejecting the statement. (24 points)

	True/Active Trail		True/Active Trail
$X_1 \perp X_2$		$X_1 \perp X_4 \mid Y_1$	
$X_1 \perp X_2 \mid Y_1$		$X_1 \perp X_4 \mid Y_3$	
$X_1 \perp X_2 \mid Y_1, Y_2$		$X_1 \perp X_4 \mid Y_3, Y_4$	
$Y_1 \perp Y_3 \mid X_2$		$X_1 \perp X_4 \mid Y_1, Y_2, Y_3, Y_4$	

Question 2 - Maximum Entropy Markov Model (40 points, 30 minutes)

A Maximum Entropy Markov Model (MEMM) has the following directed graph, where the variables $\mathbf{X_i}$ represent the states and the variables $\mathbf{O_i}$ are the observations.



A) Write down the joint distribution (up to time **t**) in terms of the CPDs (5 points).

$\mathbf{P(X_1, X_2, \ldots, X_t, O_1, O_2, \ldots, O_t) =}$

B) Notice that in most applications the observations $\mathbf{O_i}$ are always known, both at training and testing (inference) stages. Show that an MEMM can be regarded as a ***discriminative model*** by writing the following conditional distribution in terms of the CPDs. Write the full derivations. What CPDs are not used? (10 points)

$\mathbf{P(X_1, X_2, \ldots, X_t \mid O_1, O_2, \ldots, O_t) =}$

C) We intend to find $\mathbf{P(X_t \mid O_1, O_2, \ldots, O_t)}$ recursively. First, show that this distribution is readily available for **t=1** as one of the CPDs. (3 points)

D) Now, find $\mathbf{P(X_{t+1} \mid O_1, O_2, \ldots, O_{t+1})}$ in terms of the $\mathbf{P(X_t \mid O_1, O_2, \ldots, O_t)}$ and some CPDs. To do this, first write $\mathbf{P(X_{t+1} \mid O_1, O_2, \ldots, O_{t+1})}$ in terms of $\mathbf{P(X_{t+1}, X_t \mid O_1, O_2, \ldots, O_{t+1})}$. (4 points)

$\mathbf{P(X_{t+1} \mid O_1, O_2, \ldots, O_{t+1})}$ = $\mathbf{P(X_{t+1}, X_t \mid O_1, O_2, \ldots, O_{t+1})}$

E) Now break $\mathbf{P(X_{t+1}, X_t \mid O_1, O_2, \ldots, O_{t+1})}$ into the product of $\mathbf{P(X_{t+1} \mid X_t, O_1, O_2, \ldots, O_{t+1})}$ and $\mathbf{P(X_t \mid O_1, O_2, \ldots, O_{t+1})}$. Simplify the result using the conditional independence relations read from the graph. In each case mention what conditional independence relation you have used. (You may use the flow of influence, etc.). Finally, write $\mathbf{P(X_{t+1} \mid O_1, O_2, \ldots, O_{t+1})}$ in terms of $\mathbf{P(X_t \mid O_1, O_2, \ldots, O_t)}$ and one or more CPDs. (18 points)

Question 3 - Markov Random Fields (30 points, 30 minutes)

Consider an MRF in which each variables $X_{i,j}$ is indexed by two integers i and j. The joint distribution is in the form of

$$P(X_{1,1}, X_{2,1}, X_{2,2}, X_{3,1}, X_{3,2}, X_{3,3}, X_{4,1}, X_{4,2}, X_{4,3}, X_{4,4}) =$$
$$\frac{1}{Z} \prod_{i=1}^{3} \prod_{j=1}^{i} \phi_{ij}(X_{i,j}, X_{i+1,j}, X_{i+1,j+1}) \prod_{i=2}^{3} \prod_{j=1}^{i-1} \psi_{ij}(X_{i,j}, X_{i,j+1}, X_{i+1,j+1})$$

a) Draw the corresponding MRF graph. (15 points)

b) Derive $P(X_{2,2} \mid X_{1,1}, X_{2,1}, X_{3,1}, X_{3,2}, X_{3,3}, X_{4,1}, X_{4,2}, X_{4,3}, X_{4,4})$ in terms of the potentials ϕ_{ij} and ψ_{ij}. **Simplify as much as possible.** Assume that all potentials are strictly positive. (15 points)