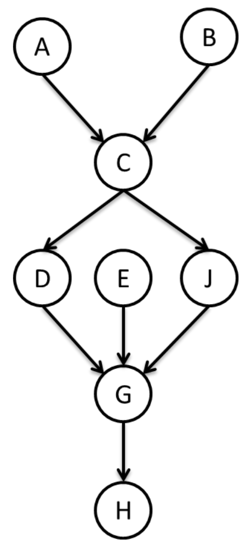
**CS 583 – Assignment 1**

## Theoretical assignments

1. We are given the random variables $X_2$, $X_3$, …, $X_n$, and $Y_2$, $Y_3$, …, $Y_\mathrm{m}$. Answer the following questions.
   a. Assuming every variable is binary, how many independent parameters are needed to represent $P(X_2, X_3, \ldots, X_n, Y_2, Y_3, \ldots, Y_m)$?
   b. Assuming every variable has three possible values, how many independent parameters are needed to represent $P(X_2, X_3, \ldots, X_n, Y_2, Y_3, \ldots, Y_m)$?
   c. Assuming each $X_i$ has $i$ possible values and similarly, every $Y_i$ has $i$ possible values, how many independent parameters are needed to represent $P(X_2, X_3, \ldots, X_n, Y_2, Y_3, \ldots, Y_m)$?
   d. Assuming every variable is binary, how many independent parameters are needed to represent $P(Y_2, Y_3, \ldots, Y_m \mid X_2, X_3, \ldots, X_n)$?
   e. Assuming every variable has three possible values, how many independent parameters are needed to represent $P(Y_2, Y_3, \ldots, Y_m \mid X_2, X_3, \ldots, X_n)$?
   f. Assuming each $X_i$ has $i$ possible values and similarly every $Y_i$ has $i$ possible values, how many independent parameters are needed to represent $P(Y_2, Y_3, \ldots, Y_m \mid X_2, X_3, \ldots, X_n)$?

**2.** We are given the following the Bayesian network. Please answer the following questions.




**a.** Write down the joint distribution as a factorization over this Bayesian network.

**b.** Assuming each variable is discrete and can take *n* possible values, how many independent parameters are needed for this Bayesian network?

**c.** Are the following independence statements true or false?

**i.** $A \perp B$

**ii.** $A \perp B \mid C$

**iii.** $A \perp B \mid J$

**iv.** $A \perp B \mid G$

**v.** $A \perp B \mid E$

**vi.** $A \perp B \mid H$

**vii.** $A \perp H$

**viii.** $A \perp H \mid J$

**ix.** $A \perp H \mid D, J$

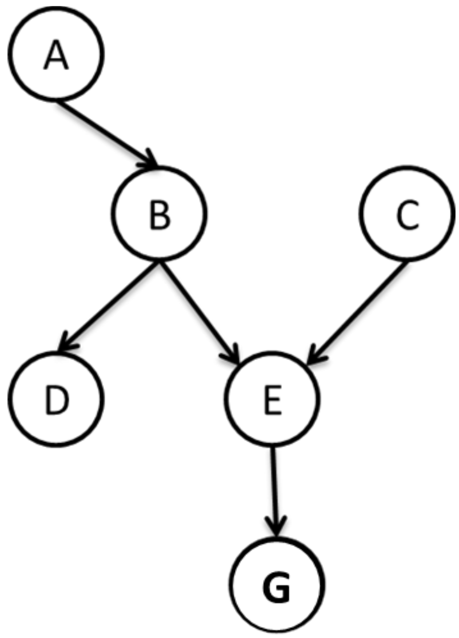
**x.** $D \perp J$

**xi.** $B \perp E$

**xii.** $B \perp E \mid J$

**xiii.** $B \perp E \mid J, H$

**3.** We have a distribution P over the variables A, B, C, D, E, and G. We would like to build a Bayesian network that is a minimal I-Map for P. In reality, you have access to P, which you can query for independencies, but for the purposes of this problem, we will assume the following structure is a P-Map for P. Create minimal I-Maps for P, using the following variable orders.




**a.** C, A, B, E, D, G

**b.** D, B, A, E, C, G

**c.** G, E, D, C, B, A

**d.** G, A, C, E, D, B

**Programming Assignments**

The code should be compatible with Python 3.7. Primary libraries: pgmpy, scikit-learn.

Submission should be either a IPython Notebook or a standalone script that provides the following output:

- Drawn structure of the Bayesian Network (if applicable)
- Inference results

**4.** We are given the following joint distribution over the random variables A, B, C, and D. Please answer the following questions. Show the necessary tables. You can (and should) share computations between different parts of the question. Usage of TableCPD is recommended

| A | B | C | D | P(A,B,C,D) |
|---|---|---|---|---|
| T | T | T | T | 0.0448 |
| T | T | T | F | 0.0252 |
| T | T | F | T | 0.0112 |
| T | T | F | F | 0.0588 |
| T | F | T | T | 0.0144 |
| T | F | T | F | 0.0144 |
| T | F | F | T | 0.0096 |
| T | F | F | F | 0.0216 |
| F | T | T | T | 0.1024 |
| F | T | T | F | 0.0576 |
| F | T | F | T | 0.0256 |
| F | T | F | F | 0.1344 |
| F | F | T | T | 0.1152 |
| F | F | T | F | 0.1152 |
| F | F | F | T | 0.0768 |
| F | F | F | F | 0.1728 |

**a.** Are A and B independent?

**b.** Are A and C independent?

**c.** Are A and C independent given B?

**d.** Are A and D independent?

**e.** Are A and D independent given C?

**5.** For the given problem, construct a Bayesian Network and general-purpose decision tree. Compare results. What differences do you observe? Explain.
Predict the probability that children would be playing outdoors given weather conditions.

| Outlook | Temp. | Humidity | Windy | Play |
|---|---|---|---|---|
| Sunny | Hot | High | False | No |
| Sunny | Hot | High | True | No |
| Overcast | Hot | High | False | Yes |
| Rainy | Mild | High | False | Yes |
| Rainy | Cool | Normal | False | Yes |
| Rainy | Cool | Normal | True | No |
| Overcast | Cool | Normal | True | Yes |
| Sunny | Mild | High | False | No |
| Sunny | Cool | Normal | False | Yes |
| Rainy | Mild | Normal | False | Yes |
| Sunny | Mild | Normal | True | Yes |
| Overcast | Mild | High | True | Yes |
| Overcast | Hot | Normal | False | Yes |
| Rainy | Mild | High | True | No |