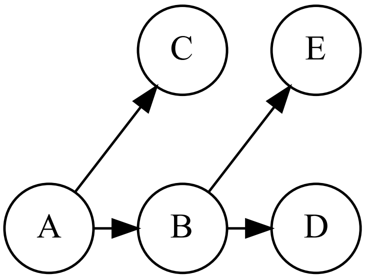
Assignment 6: Bayesian Networks — Name: *Niels Van den Broeck*
Deadline: December 3, 2023 (18 points) — Student number: *s0203844*

Probability (2 points)



	P(A)
$+a$	0.25
$-a$	0.75

P$(B \mid A)$	$+b$	$-b$
$+a$	0.5	0.5
$-a$	0.25	0.75

P$(C \mid A)$	$+c$	$-c$
$+a$	0.2	0.8
$-a$	0.6	0.4

P$(D \mid B)$	$+d$	$-d$
$+b$	0.6	0.4
$-b$	0.8	0.2

P$(E \mid B)$	$+e$	$-e$
$+b$	0.25	0.75
$-b$	0.1	0.9

1. Using the Bayes net and conditional probability tables above, calculate the following quantities: (2 pts)
   (a) P$(+b \mid +a) =$ *0.5*
   (b) P$(+a, +b) =$ *0.125*
   (c) P$(+a \mid +b) =$ *0.4*
   (d) P$(-e, +a) =$ *0.825*
   (e) P$(D \mid A) =$

P$(D \mid A)$	$+d$	$-d$
$+a$	*0.7*	*0.3*
$-a$	*0.75*	*0.25*

Independence (8 points)

2. For each of the following equations, select the minimal set of conditional independence assumptions necessary for the equation to be true. (4 pts)

   (a) P$(A, C) =$ P$(A \mid B)$P(C)

   √ $A \perp\!\!\!\perp B$ ☐ $A \perp\!\!\!\perp B \mid C$ √ $A \perp\!\!\!\perp C$ ☐ $A \perp\!\!\!\perp C \mid B$
   ☐ $B \perp\!\!\!\perp C$ ☐ $B \perp\!\!\!\perp C \mid A$ ☐ No independence assumptions needed

   (b) P$(A \mid B, C) = \frac{\text{P}(A)\text{P}(B|A)\text{P}(C|A)}{\text{P}(B|C)\text{P}(C)}$

   ☐ $A \perp\!\!\!\perp B$ ☐ $A \perp\!\!\!\perp B \mid C$ ☐ $A \perp\!\!\!\perp C$ ☐ $A \perp\!\!\!\perp C \mid B$
   ☐ $B \perp\!\!\!\perp C$ √ $B \perp\!\!\!\perp C \mid A$ ☐ No independence assumptions needed

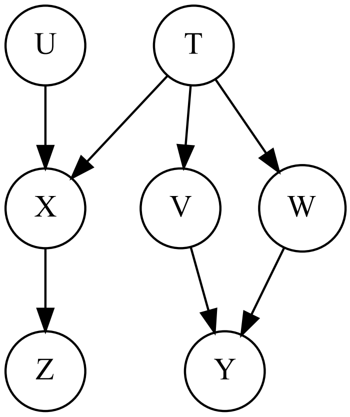
   (c) P$(A, B) = \sum_c$ P$(A \mid B, c)$P$(B|c)$P(c)

   ☐ $A \perp\!\!\!\perp B$ ☐ $A \perp\!\!\!\perp B \mid C$ ☐ $A \perp\!\!\!\perp C$ ☐ $A \perp\!\!\!\perp C \mid B$
   ☐ $B \perp\!\!\!\perp C$ ☐ $B \perp\!\!\!\perp C \mid A$ √ *No independence assumptions needed*

   (d) P$(A, B \mid C, D) =$ P$(A \mid C, D)$P$(B \mid A, C, D)$

   ☐ $A \perp\!\!\!\perp B$ ☐ $A \perp\!\!\!\perp B \mid C$ ☐ $A \perp\!\!\!\perp C$ ☐ $A \perp\!\!\!\perp C \mid B$
   ☐ $B \perp\!\!\!\perp C$ ☐ $B \perp\!\!\!\perp C \mid A$ √ *No independence assumptions needed*

3. Indicate whether each of the following conditional independence relationships is guaranteed to be true in the Bayes Net below. If the independence relationship does not hold, identify all active (d-connected) paths in the graph. (4 pts)



(a) $T \perp\!\!\!\perp Y$
Independence is ○ Guaranteed √ *Not Guaranteed*
$T \rightarrow W \rightarrow Y$ and $T \rightarrow V \rightarrow Y$ are active triples

(b) $T \perp\!\!\!\perp Y \mid W$
Independence is ○ Guaranteed √ *Not Guaranteed*
$T \rightarrow W \rightarrow Y$ is now inactive, but $T \rightarrow V \rightarrow Y$ still is active.

(c) $U \perp\!\!\!\perp T$
Independence is √ *Guaranteed* ○ Not Guaranteed

(d) $U \perp\!\!\!\perp T \mid Z$
Independence is ○ Guaranteed √ *Not Guaranteed*
In the V-structure: $U \rightarrow X \leftarrow T$, the descendent of X, Z is observed, which makes it an active path.

(e) $Z \perp\!\!\!\perp U$
Independence is ○ Guaranteed √ *Not Guaranteed*
the path: $U \rightarrow X \rightarrow Z$ is an active path.

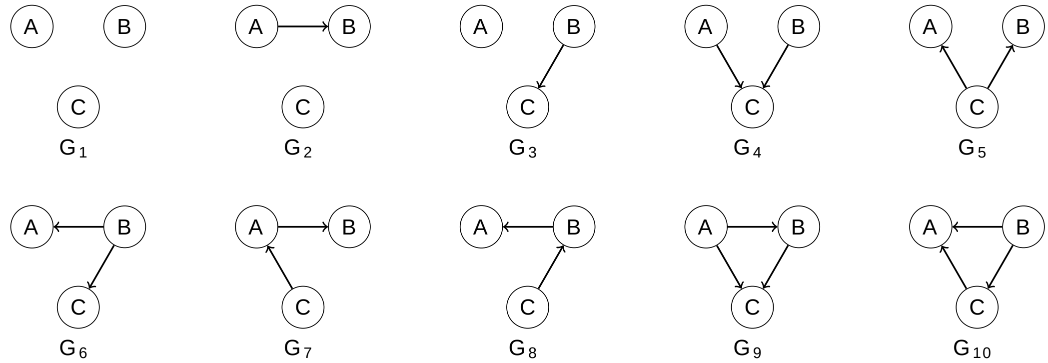
(f) $Z \perp\!\!\!\perp Y \mid V$
Independence is ○ Guaranteed √ *Not Guaranteed*
$Z \leftarrow X \leftarrow T \rightarrow W \rightarrow Y$ is an active path which makes it not guaranteed.

(g) $Z \perp\!\!\!\perp Y \mid T, W$
Independence is √ *Guaranteed* ○ Not Guaranteed

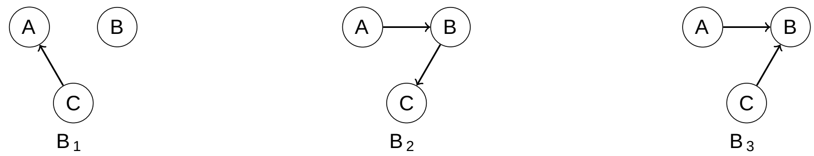
(h) $Z \perp\!\!\!\perp W$
Independence is ○ Guaranteed √ *Not Guaranteed*
same reason as (f), $Z \leftarrow X \leftarrow T \rightarrow W$ is an active path.

Representation (4 points)

4. We are given the following ten Bayes nets, labeled $\mathbf{G}_1$ to $\mathbf{G}_{10}$: (4 pts)



and the following three Bayes nets, labeled $\mathbf{B}_1$ to $\mathbf{B}_3$:



(a) Assume we know that a joint distribution d_1 (over A, B, C) can be represented by Bayes net $\mathbf{B}_1$. Mark all of the following Bayes nets that are guaranteed to be able to represent d_1.

☐ G_1 ☐ G_2 ☐ G_3 √ G_4 √ G_5
☐ G_6 √ G_7 ☐ G_8 √ G_9 √ G_{10}
☐ None of the above

(b) Assume we know that a joint distribution d_2 (over A, B, C) can be represented by Bayes net $\mathbf{B}_2$. Mark all of the following Bayes nets that are guaranteed to be able to represent d_2.

☐ G_1 ☐ G_2 ☐ G_3 ☐ G_4 ☐ G_5
√ G_6 ☐ G_7 √ G_8 √ G_9 √ G_{10}
☐ None of the above

(c) Assume we know that a joint distribution d_3 (over A, B, C) **cannot** be represented by Bayes net $\mathbf{B}_3$. Mark all of the following Bayes nets that are guaranteed to be able to represent d_3.

☐ G_1 ☐ G_2 ☐ G_3 ☐ G_4 ☐ G_5
☐ G_6 ☐ G_7 ☐ G_8 √ G_9 √ G_{10}
☐ None of the above

(d) Assume we know that a joint distribution d_4 (over A, B, C) can be represented by Bayes nets $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$. Mark all of the following Bayes nets that are guaranteed to be able to represent d_4.

√ G_1 √ G_2 √ G_3 √ G_4 √ G_5
√ G_6 √ G_7 √ G_8 √ G_9 √ G_{10}
☐ None of the above

Inference (4 points)

5. Using the same Bayes Net from question 3, we want to compute $\mathsf{P}(Y \mid +z)$. All variables have binary domains. Assume we run variable elimination to compute the answer to this query, with the following variable elimination ordering: U, T, X, V, W. (4 pts)

   Complete the following description of the factors generated in this process:

   (a) After inserting evidence, we have the following factors to start out with:
   $\mathsf{P}(U), \mathsf{P}(T), \mathsf{P}(X \mid U, T), \mathsf{P}(V \mid T), \mathsf{P}(W \mid T), \mathsf{P}(+z \mid X), \mathsf{P}(Y \mid V, W)$

   (b) When eliminating U we generate a new factor f_1 as follows, which leaves us with the factors: $f_1(X, T) = \sum_u \mathsf{P}(u)\mathsf{P}(X \mid u, T)$
   Factors: $\mathsf{P}(T), \mathsf{P}(V \mid T), \mathsf{P}(W \mid T), \mathsf{P}(+z \mid X), \mathsf{P}(Y \mid V, W), f_1(X, T)$

   (c) When eliminating T we generate a new factor f_2 as follows, which leaves us with the factors:
   $f_2(X, V, W) = \sum_t \mathsf{P}(t) f_1(X, t)\mathsf{P}(V \mid t)\mathsf{P}(W \mid t)$
   Factors: $\mathsf{P}(+z \mid X), \mathsf{P}(Y \mid V, W), f_2(X, V, W)$

   (d) When eliminating X we generate a new factor f_3 as follows, which leaves us with the factors:
   $f_3(+z, V, W) = \sum_x \mathsf{P}(+z \mid x) f_2(x, V, W)$
   Factors: $\mathsf{P}(Y \mid V, W), f_3(+z, V, W)$

   (e) When eliminating V we generate a new factor f_4 as follows, which leaves us with the factors:
   $f_4(+z, Y, W) = \sum_v \mathsf{P}(Y \mid v, W) f_3(+z, v, W)$
   Factors: $f_4(+z, Y, W)$

   (f) When eliminating W we generate a new factor f_5 as follows, which leaves us with the factors:
   $f_5(+z, Y) = \sum_w f_4(+z, Y, w)$
   Factors: $f_5(+z, Y)$

   (g) How would you obtain $\mathsf{P}(Y \mid +z)$ from the factors left above?
   The only factor left is $f_5(+z, Y)$. $\mathsf{P}(Y \mid +z)$ *Can be obtained by normalizing this factor: The normalizer Z is* $\sum_{y'} f_5(+z, y')$. $\mathsf{P}(Y \mid +z) = \frac{1}{Z} f_5(+z, Y) = \frac{f_5(+z,Y)}{\sum_{y'} f_5(+z,y')}$.

   (h) What is the size of the largest factor that gets generated during the above process?
   $f_2(X, V, W)$ *has 3 unconditioned variables which means its size is* $2^3 = 8$.

   (i) Does there exist a better elimination ordering (one which generates smaller largest factors)? Argue why not or give an example.
   X,U,T,V,W or X,U,T,W,V would be a better elimination order because it would only have factors of maximum size of $2^2 = 4$. *Taking X first generates the factor* $f_1(+z, T, U)$ *with factors:* $P(U), P(T), f_1(+z, T, U), P(V \mid T), P(W \mid T), P(Y \mid V, W)$. *Now we eliminate*

U which creates $f_2(T, +z)$ *with factors:* $P(T), f_2(T, +z), P(V \mid T), P(W \mid T), P(Y \mid V, W)$. *Then T is chosen which generates* $f_3(+z, V, W)$ *with factors* $f_3(+z, V, W), P(Y \mid V, W)$. *For the remaining variables, the ordering doesn't matter because they will not exceed the maximum size of* $2^2 = 4$.