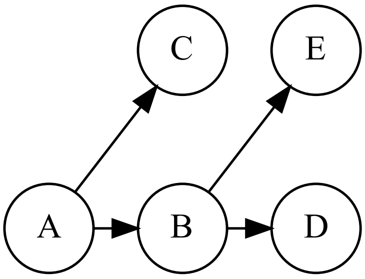
Assignment 6: Bayesian Networks Name: *Orfeo Terkuçi*
Deadline: December 3, 2023 (18 points) Student number: *s0213863*

Probability (2 points)



	$\mathsf{P}(A)$
$+a$	0.25
$-a$	0.75

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

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

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

$\mathsf{P}(E \mid B)$	$+e$	$-e$
$+b$	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) $\mathsf{P}(+b \mid +a) =$ *0.5*
   (b) $\mathsf{P}(+a, +b) = \mathsf{P}(+a)\mathsf{P}(+b \mid +a) = 0.25 \cdot 0.5 = 0.125$
   (c) $\mathsf{P}(+a \mid +b) = \dfrac{\mathsf{P}(+b \mid +a) \cdot \mathsf{P}(+a)}{\mathsf{P}(+b)} = \dfrac{0.5 \cdot 0.25}{0.5 \cdot 0.25 + 0.25 \cdot 0.75} = \dfrac{0.125}{0.125 + 0.1875} = \dfrac{0.125}{0.3125} = 0.4$
   (d) $\mathsf{P}(-e, +a) = \mathsf{P}(+a)\mathsf{P}(-e \mid +a) = 0.25 \cdot (0.75 \cdot 0.5 + 0.9 \cdot 0.5) = 0.25 \cdot 0.825 = 0.20625$
   (e) $\mathsf{P}(D \mid A) =$

$\mathsf{P}(D \mid A)$	$+d$	$-d$
$+a$	*0.175*	*0.075*
$-a$	*0.5625*	*0.1875*

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) $\mathsf{P}(A, C) = \mathsf{P}(A \mid B)\mathsf{P}(C)$

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

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

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

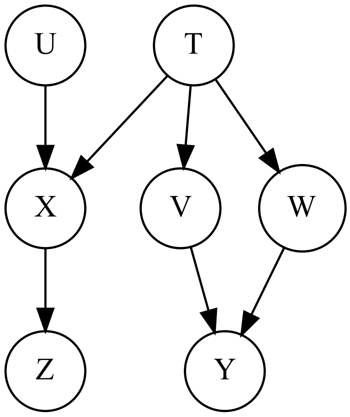
   (c) $\mathsf{P}(A, B) = \sum_c \mathsf{P}(A \mid B, c)\mathsf{P}(B|c)\mathsf{P}(c)$

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

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

   ☐ $A \perp\!\!\!\perp B$ ☐ $A \perp\!\!\!\perp B \mid C$ ☐ $A \perp\!\!\!\perp C$ ☐ $A \perp\!\!\!\perp C \mid B$
   ☐ $B \perp\!\!\!\perp C$ ☐ $B \perp\!\!\!\perp C \mid A$ √ *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*
$T \to V \to Y$
$T \to W \to Y$

(b) $T \perp\!\!\!\perp Y \mid W$
Independence is ○ Guaranteed √ *Not Guaranteed*
$T \to V \to Y$

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

(d) $U \perp\!\!\!\perp T \mid Z$
Independence is ○ Guaranteed √ *Not Guaranteed*
$(U, T) \to X \to Z(shaded)$

(e) $Z \perp\!\!\!\perp U$
Independence is ○ Guaranteed √ *Not Guaranteed*
$U \to X \to Z$

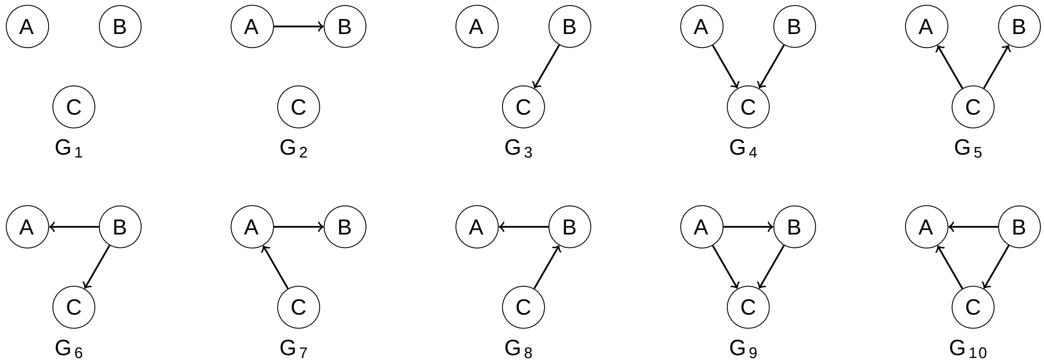
(f) $Z \perp\!\!\!\perp Y \mid V$
Independence is ○ Guaranteed √ *Not Guaranteed*
$T \to W \to Y$
$T \to X \to Z$
$T \to (X, W)$

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

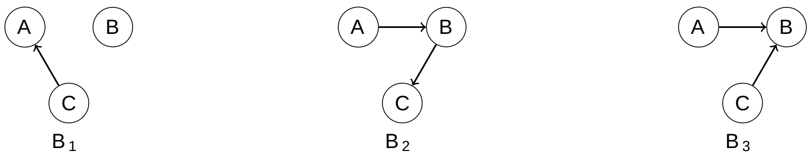
(h) $Z \perp\!\!\!\perp W$
Independence is ○ Guaranteed √ *Not Guaranteed*
$T \to (X, W)$
$T \to X \to Z$

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:
   $f_2(V, W, X) = \sum_t \mathsf{P}(t)\mathsf{P}(V \mid t)\mathsf{P}(W, t) f_1(X, t)$
   Factors: $\mathsf{P}(+z, X), \mathsf{P}(Y \mid V, W), f_2(V, W, X)$

   (d) When eliminating X we generate a new factor f_3 as follows, which leaves us with the factors:
   $f_3(+z, V, W) = \sum_x \mathsf{P}(+z, x) f_2(V, W, x)$
   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: $f_4(Y, +z, W)$

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

   (g) How would you obtain $\mathsf{P}(Y \mid +z)$ from the factors left above?
   $$\mathsf{P}(Y \mid +z) = \frac{f_5(Y, +z)}{\sum_{y_1} f_5(+z, y_1)}$$

   (h) What is the size of the largest factor that gets generated during the above process?
   The largest factor is $f_2(V, W, X)$. *It has three unconditional variables and each of them has two possible values, so the factor has* $2^3 = 8$ *probability entries.*

   (i) Does there exist a better elimination ordering (one which generates smaller largest factors)? Argue why not or give an example.
   Yes. One optimal ordering is U, X, T, V, W.
   From this ordering all the factors are of size 2 and we get a total of $2^2 = 4$ *probability entries.*