Assignments of the subject "Probabilistic Methods"

Topic 1. Bayesian Network Fundamentals

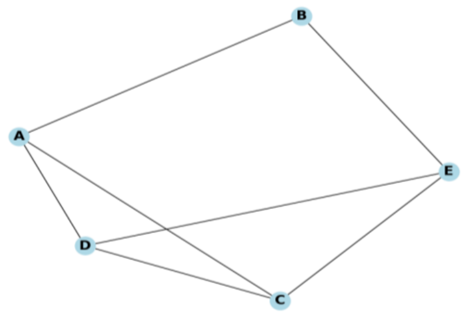
Exercise 1.1.

Let be an undirected graph G containing five nodes and the following links: A–B, A–C, A–D, C–D, B–E, C–E, and D–E.

1. For each of the following relationships, say whether it is true or false, listing all the paths between the first two variables and pointing out which are active and which are blocked.

   a) $I_G(A, B)$
   b) $I_G(A, D)$
   c) $I_G(B, C)$
   d) $I_G(C, D)$
   e) $I_G(A, E)$
   f) $I_G(B, C|A)$
   g) $I_G(B, D|C)$
   h) $I_G(B, C|A, E)$
   i) $I_G(B, D|A, E)$
   j) $I_G(A, E|B, C)$
   k) $I_G(A, E|B, C, D)$
   l) $I_G(B, C|A, D, E)$

Solution 1.1.



IG(A, B): False
Paths:
- A → B: Active
- A → C → D → E → B: Blocked
- A → C → E → B: Blocked
- A → D → C → E → B: Blocked
- A → D → E → B: Blocked

IG(A, D): False
Paths:
- A → B → E → C → D: Blocked
- A → B → E → D: Blocked
- A → C → D: Blocked
- A → C → E → D: Blocked
- A → D: Active

IG(B, C): True
Paths:
- B → A → C: Blocked
- B → A → D → C: Blocked
- B → A → D → E → C: Blocked
- B → E → C: Blocked
- B → E → D → A → C: Blocked
- B → E → D → C: Blocked

IG(C, D): False
Paths:
- C → A → B → E → D: Blocked
- C → A → D: Blocked
- C → D: Active
- C → E → B → A → D: Blocked
- C → E → D: Blocked

IG(A, E): True
Paths:
- A → B → E: Blocked
- A → C → D → E: Blocked
- A → C → E: Blocked
- A → D → C → E: Blocked
- A → D → E: Blocked

IG(B, C | A): False
Paths:
- B → A → C: Active

- B → A → D → C: Blocked
- B → A → D → E → C: Blocked
- B → E → C: Blocked
- B → E → D → A → C: Blocked
- B → E → D → C: Blocked

IG(B, D | C): True
Paths:
- B → A → C → D: Blocked
- B → A → C → E → D: Blocked
- B → A → D: Blocked
- B → E → C → A → D: Blocked
- B → E → C → D: Blocked
- B → E → D: Blocked

IG(B, C | A, E): False
Paths:
- B → A → C: Active
- B → A → D → C: Blocked
- B → A → D → E → C: Blocked
- B → E → C: Active
- B → E → D → A → C: Blocked
- B → E → D → C: Blocked

IG(B, D | A, E): False
Paths:
- B → A → C → D: Blocked
- B → A → C → E → D: Blocked
- B → A → D: Active
- B → E → C → A → D: Blocked
- B → E → C → D: Blocked
- B → E → D: Active

IG(A, E | B, C): False
Paths:
- A → B → E: Active
- A → C → D → E: Blocked
- A → C → E: Active
- A → D → C → E: Blocked
- A → D → E: Blocked

IG(A, E | B, C, D): False
Paths:
- A → B → E: Active
- A → C → D → E: Active
- A → C → E: Active
- A → D → C → E: Active
- A → D → E: Active

IG(B, C | A, D, E): False
Paths:
- B → A → C: Active
- B → A → D → C: Active
- B → A → D → E → C: Active
- B → E → C: Active
- B → E → D → A → C: Active
- B → E → D → C: Active

Exercise 1.2.

Let be a directed graph G containing 7 nodes and the following links: $A \longrightarrow C$, $B \longrightarrow C$, $B \longrightarrow D$, $C \longrightarrow E$, $C \longrightarrow F$, $D \longrightarrow F$, $E \longrightarrow G$ and $F \longrightarrow G$.

1. For each of the following relationships, say whether it is true or false, listing all the paths between the first two variables and pointing out which are active and which are blocked.

   a) $IG(A, B)$

   b) $IG(A, C)$

   c) $IG(A, D)$

   d) $IG(A, E)$

   e) $IG(A, G)$

   f) $IG(D, E)$

   g) $GI(C, G)$.

2. For each of the following relationships, say whether it is true or false, listing all the paths between the first two variables and indicating for each whether it is active, inactive, or blocked.

   a) $I_G(A, B|C)$

   b) $I_G(A, B|E)$

   c) $I_G(A, B|F)$

   d) $I_G(A, B|G)$

   e) $I_G(A, D|C)$

   f) $I_G(A, D|E)$

   g) $I_G(A, F|C)$

   h) $I_G(A, F|E)$

   i) $I_G(A, F|G)$

   j) $I_G(A, E|F)$

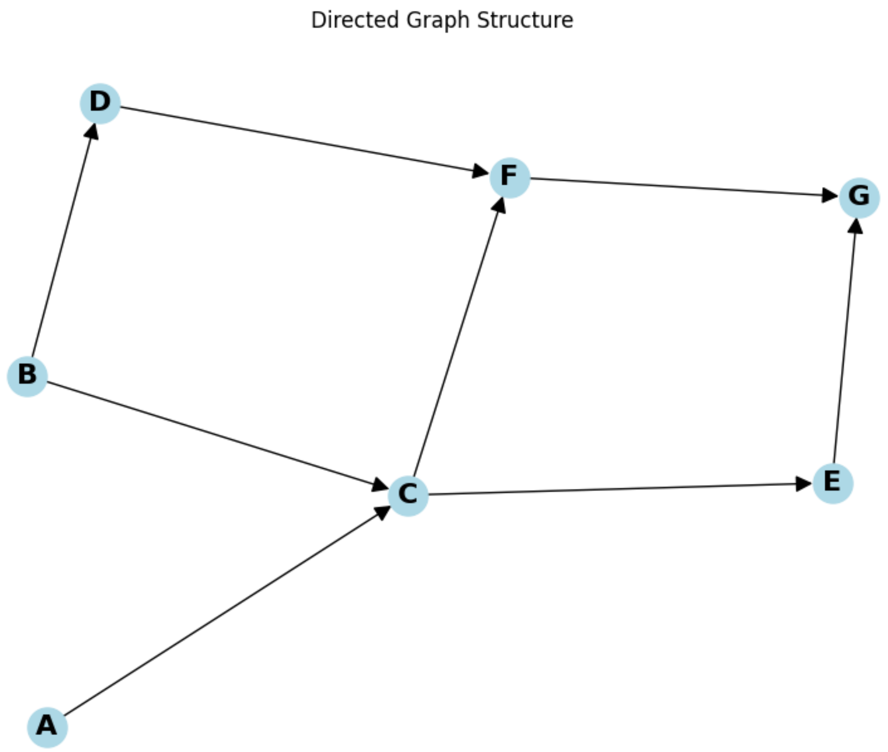
   k) $I_G(B, E|C, D)$

   l) $I_G(B, G|D, E)$

   m) $I_G(D, E|C, G)$

   n) $I_G(B, F|A, D, G)$.

Solution 1.2.



<u>Part 1: Unconditional Independence Relationships</u>

a) IG(A, B): True
Paths:
• A → C ← B: Blocked (C is a collider)
• A → C → E → G ← F ← D ← B: Blocked (G is a collider)
• A → C → F ← D ← B: Blocked (F is a collider)
Conclusion: All paths are blocked.

b) IG(A, C): False
Paths:
• A → C: Active
Conclusion: Active path exists.

c) IG(A, D): True
Paths:
• A → C ← B → D: Blocked (C is a collider)
• A → C → E → G ← F ← D: Blocked (G is a collider)
• A → C → F ← D: Blocked (F is a collider)
Conclusion: All paths are blocked.

d) IG(A, E): False
Paths:
• A → C ← B → D → F → G ← E: Blocked (C is a collider)
• A → C → E: Active (C is not conditioned on)
• A → C → F → G ← E: Blocked (G is a collider)
Conclusion: Active path exists.

e) IG(A, G): False
Paths:
• A → C ← B → D → F → G: Blocked (C is a collider)
• A → C → E → G: Active (C is not conditioned on)
• A → C → F → G: Active (C is not conditioned on)
Conclusion: Active path exists.

f) IG(D, E): True
Paths:
• D ← B → C → E: Blocked (B is a collider)
• D ← B → C → F → G ← E: Blocked (B is a collider)
• D → F ← C → E: Blocked (F is a collider)
• D → F → G ← E: Blocked (G is a collider)
Conclusion: All paths are blocked.

g) IG(C, G): False
Paths:
• C ← B → D → F → G: Blocked (B is a collider)
• C → E → G: Active (E is not conditioned on)
• C → F → G: Active (F is not conditioned on)
Conclusion: Active path exists.

Part 2: Conditional Independence Relationships

a) IG(A, B | C): False
Paths:
• A → C → F ← D ← B: Blocked (C is conditioned on)
• A → C → E → G ← F ← D ← B: Blocked (C is conditioned on)
• A → C ← B: Active (C is an observed collider)
Conclusion: Active path exists.

b) IG(A, B | E): True
Paths:
• A → C → F ← D ← B: Blocked (F is a collider)
• A → C → E → G ← F ← D ← B: Blocked (E is conditioned on)
• A → C ← B: Blocked (C is a collider)
Conclusion: All paths are blocked.

c) IG(A, B | F): False
Paths:
• A → C → F ← D ← B: Active (C is not conditioned on)
• A → C → E → G ← F ← D ← B: Blocked (G is a collider)
• A → C ← B: Blocked (C is a collider)
Conclusion: Active path exists.

d) IG(A, B | G): False
Paths:
• A → C → F ← D ← B: Blocked (F is a collider)
• A → C → E → G ← F ← D ← B: Active (C is not conditioned on)
• A → C ← B: Blocked (C is a collider)
Conclusion: Active path exists.

e) IG(A, D | C): True
Paths:
• A → C → F ← D: Blocked (C is conditioned on)
• A → C → E → G ← F ← D: Blocked (C is conditioned on)
• A → C ← B → D: Blocked (B is a collider)
Conclusion: All paths are blocked.

f) IG(A, D | E): True
Paths:
• A → C → F ← D: Blocked (F is a collider)
• A → C → E → G ← F ← D: Blocked (E is conditioned on)
• A → C ← B → D: Blocked (C is a collider)
Conclusion: All paths are blocked.

g) IG(A, F | C): True
Paths:
• A → C → F: Blocked (C is conditioned on)
• A → C → E → G ← F: Blocked (C is conditioned on)
• A → C ← B → D → F: Blocked (B is a collider)
Conclusion: All paths are blocked.

h) IG(A, F | E): False
Paths:
• A → C → F: Active (C is not conditioned on)
• A → C → E → G ← F: Blocked (E is conditioned on)
• A → C ← B → D → F: Blocked (C is a collider)
Conclusion: Active path exists.

i) IG(A, F | G): False
Paths:
• A → C → F: Active (C is not conditioned on)
• A → C → E → G ← F: Active (C is not conditioned on)
• A → C ← B → D → F: Blocked (C is a collider)
Conclusion: Active path exists.

j) IG(A, E | F): False
Paths:
• A → C → F → G ← E: Blocked (F is conditioned on)
• A → C → E: Active (C is not conditioned on)
• A → C ← B → D → F → G ← E: Blocked (C is a collider)
Conclusion: Active path exists.

k) IG(B, E | C, D): True
Paths:
• B → C → F → G ← E: Blocked (C is conditioned on)
• B → C → E: Blocked (C is conditioned on)
• B → D → F ← C → E: Blocked (D is conditioned on)
• B → D → F → G ← E: Blocked (D is conditioned on)
Conclusion: All paths are blocked.

l) IG(B, G | D, E): False
Paths:
• B → C → F → G: Active (C is not conditioned on)
• B → C → E → G: Blocked (E is conditioned on)
• B → D → F ← C → E → G: Blocked (D is conditioned on)
• B → D → F → G: Blocked (D is conditioned on)
Conclusion: Active path exists.

m) IG(D, E | C, G): False
Paths:
• D → F ← C → E: Blocked (F is a collider)
• D → F → G ← E: Active (F is not conditioned on)
• D ← B → C → F → G ← E: Blocked (B is a collider)
• D ← B → C → E: Blocked (B is a collider)
Conclusion: Active path exists.

n) IG(B, F | A, D, G): False
Paths:
• B → C → F: Active (C is not conditioned on)
• B → C → E → G ← F: Active (C is not conditioned on)
• B → D → F: Blocked (D is conditioned on)
Conclusion: Active path exists.

Exercise 1.3.

The probability distribution P is given by the following table.

to	b	c	P (a, b, c)
+a	+b	+c	0'112
+a	+b	$\neg c$	0'048
+a	$\neg b$	+c	0'064
+a	$\neg b$	$\neg c$	0'576
$\neg a$	+b	+c	0'056
$\neg a$	+b	$\neg c$	0'024
$\neg a$	$\neg b$	+c	0'012
$\neg a$	$\neg b$	$\neg c$	0'108

1. Point out which of these relationships are true and which are false. Note that, as stated on page 32 of [1], $IP\,(X, Y\,)$ means that X and Y are independent variables (cf. Def. 1.19 of [1]) and $IP\,(X, Y\,Z)$ means that they are independent given Z (cf. Def. 1.22). For the first and fifth, indicate in detail the calculations you have made.

   a) $IP\,(A, B)$

   b) $IP\,(A, C)$

   c) $IP\,(B, C)$

   d) $IP\,(B, C|A)$

   e) $IP\,(A, C|B)$

   f) $IP\,(A, B|C)$

2. Draw **all** undirected graphs that are independence maps (*I-maps*) of P . *Tip:* The graph must have one node for each variable of P . Pay attention, because there are more graphs than one can think at first.

3. Draw **all** acyclic directed graphs that are independence maps (I-maps) of P .

Solution 1.3.

1)

Independence Relationships

IP(A, B): False

IP(A, C): False

IP(B, C): False

IP(B, C | A): False

IP(A, C | B): True

IP(A, B | C): False

Detailed Calculations

IP(A, B): False

Calculations:

- P(A=True, B=True) = 0.160 P(A=True) × P(B=True) = 0.800 × 0.240 = 0.192
- P(A=True, B=False) = 0.640 P(A=True) × P(B=False) = 0.800 × 0.760 = 0.608
- P(A=False, B=True) = 0.080 P(A=False) × P(B=True) = 0.200 × 0.240 = 0.048
- P(A=False, B=False) = 0.120 P(A=False) × P(B=False) = 0.200 × 0.760 = 0.152

Conclusion: A and B are not independent because the joint probabilities differ from the product of the marginals.

IP(A, C | B): True

Given B=True

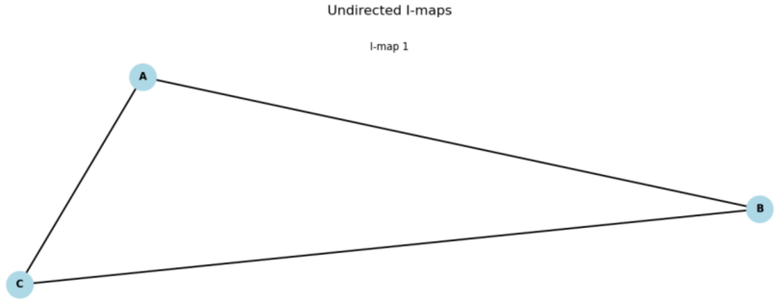
- P(A=True, C=True | B=True) = 0.467
P(A=True | B=True) × P(C=True | B=True) = 0.667 × 0.700 = 0.467
- P(A=True, C=False | B=True) = 0.200
P(A=True | B=True) × P(C=False | B=True) = 0.667 × 0.300 = 0.200
- P(A=False, C=True | B=True) = 0.233
P(A=False | B=True) × P(C=True | B=True) = 0.333 × 0.700 = 0.233
- P(A=False, C=False | B=True) = 0.100
P(A=False | B=True) × P(C=False | B=True) = 0.333 × 0.300 = 0.100

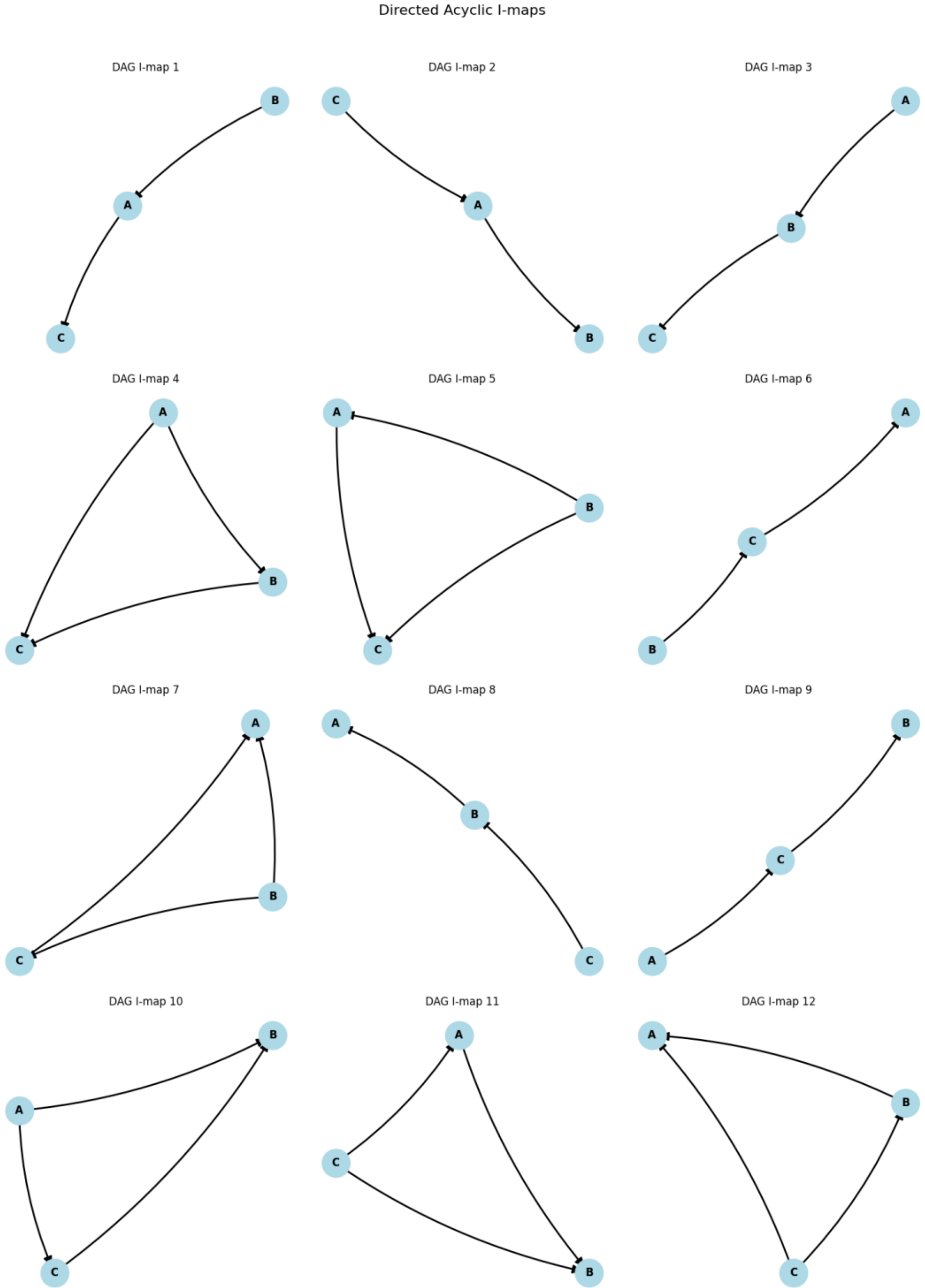
Given B=False

- P(A=True, C=True | B=False) = 0.084
P(A=True | B=False) × P(C=True | B=False) = 0.842 × 0.100 = 0.084
- P(A=True, C=False | B=False) = 0.758
P(A=True | B=False) × P(C=False | B=False) = 0.842 × 0.900 = 0.758
- P(A=False, C=True | B=False) = 0.016
P(A=False | B=False) × P(C=True | B=False) = 0.158 × 0.100 = 0.016
- P(A=False, C=False | B=False) = 0.142
P(A=False | B=False) × P(C=False | B=False) = 0.158 × 0.900 = 0.142

Conclusion: A and C are conditionally independent given B because all joint conditionals equal the product of the individual conditionals.

2 & 3)


Undirected I-maps
I-map 1
A
B
C


Directed Acyclic I-maps
DAG I-map 1
DAG I-map 2
DAG I-map 3
DAG I-map 4
DAG I-map 5
DAG I-map 6
DAG I-map 7
DAG I-map 8
DAG I-map 9
DAG I-map 10
DAG I-map 11
DAG I-map 12

Exercise 1.4.

In one country, the prevalence of typhoid fever is 0.001 and that of tuberculosis is 0.01. Typhoid fever always causes fever, and bradycardia (slow heartbeat) in 40% of cases. TB causes fever in 60% of cases and tachycardia (faster-than-normal heart rate) in 58%. The prevalence of fever in patients who do not suffer from either of these two diseases is 1.5%, bradycardia 0.05% and tachycardia 1.3%.

1. Draw the graph of a Bayesian network for this problem. (In the next exercise you will have to draw the graph for the naïve Bayesian method.)

2. Indicate which values each variable takes.

3. Indicate the conditional probabilities, in the form of tables, that define the model. You may find it helpful to apply the OR gate, which is explained in section 3.2 of [1].

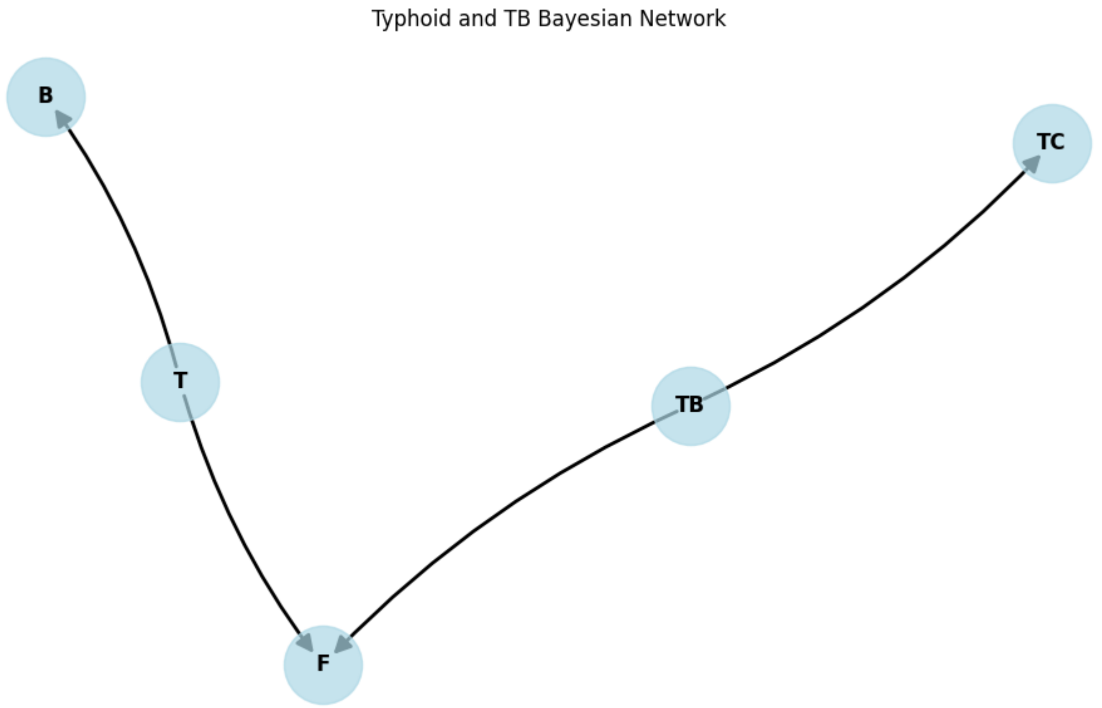
4. Point out the hypotheses you are using to solve this problem and discuss whether they are reasonable or not, i.e., whether they seem to be a good approximation.

5. Indicate in a table what the diagnosis (the probability of each disease) is for each of the possible combinations of findings: fever, non-fever, tachycardia, bradycardia, normal rhythm, fever and tachycardia, fever and normal rhythm, etc.? If you want, you can use the OpenMarkov program to complete the table, but in that case you must perform "by hand" and show the detailed calculations for two of those combinations.

6. Experience shows that, when tuberculosis is present, fever and tachycardia are associated in most cases (i.e., tuberculosis usually produces tachycardia if and only if it produces fever). Does this observation call into question the validity of the results obtained in the previous section?

Solution 1.4.

1. Bayesian Network Structure



2. Variable Values

All variables are binary:

Variable	**Meaning**	**Values**
T	Typhoid fever	True/False
TB	Tuberculosis	True/False
F	Fever	True/False
B	Bradycardia	True/False
TC	Tachycardia	True/False

3. Conditional Probability Tables (CPTs)

CPT for TB

TB	**P(TB)**
True	0.0100
False	0.9900

CPT for T

T	**P(T)**
True	0.0010
False	0.9990

CPT for F (given T and TB)

T	**TB**	**P(F=True)**	**P(F=False)**
False	False	0.0150	0.9850
True	False	1.0000	0.0000
False	True	0.6000	0.4000
True	True	1.0000	0.0000

CPT for TC (given TB)

TB	**P(TC=True)**	**P(TC=False)**
False	0.0130	0.9870
True	0.5800	0.4200

CPT for B (given T)

T	**P(B=True)**	**P(B=False)**
False	0.0005	0.9995
True	0.4000	0.6000

4. Assumptions and Reasonableness

Assumptions:

1. **Independence**: T and TB are marginally independent (no known medical correlation).
2. **Conditional Independence**: Given diseases, symptoms (F, Br, Ta) are independent of each other.
3. **OR Gate**: Fever results from an OR of T and TB (with respective probabilities).
4. **Baseline Symptoms**: When both T and TB are false, use general population data.

Reasonableness:

- Assumption 1 is fine due to low prevalence.
- Assumption 2 is a simplification, but common in probabilistic modeling.
- Assumption 3 (OR gate for fever) makes sense, though it's a simplification.
- Assumption 4 is practical but could be inaccurate if other causes exist.

However, as noted in **point 6**, real-world correlation (e.g., TB causing fever and tachycardia together) may challenge the independence assumption. It might be better to model **F** and **Ta** as jointly dependent on **TB** (introduce a hidden common cause or make F a parent of Ta).

5. Diagnosis Table (Example Calculations)

Example 1: Patient has Fever and Tachycardia

Diagnosing Typhoid (T)

- Calculating P(T=True | F=True, TC=True):

Joint probability components:

- P(T=True) = 0.0010
- P(F=True | T=True, TB=True) = 2.0000
- P(TC=True | TB=True) = 0.5930

Final joint probability = 0.001186

Normalizing constant:

- P(F=True | T=True, TB=True) = 2.6150
- P(TC=True | TB=True) = 0.5930
  Final = 1.550695

P(T=True | F=True, TC=True) = 0.000765

- Calculating P(T=False | F=True, TC=True):

Joint probability components:

  - P(T=False) = 0.9990
  - P(F=True | T=False, TB=True) = 0.6150
  - P(TC=True | TB=True) = 0.5930

Final joint probability = 0.364330

Normalizing constant (same as above) = 1.550695

P(T=False | F=True, TC=True) = 0.234946

Final probabilities for T given F=True, TC=True:

- P(T=True | F=True, TC=True) = 0.003245
- P(T=False | F=True, TC=True) = 0.996755

Diagnosing Tuberculosis (TB)

- Calculating P(TB=True | F=True, TC=True):

Joint probability components:

  - P(TB=True) = 0.0100
  - P(F=True | TB=True, T=True) = 1.6000
  - P(TC=True | TB=True) = 0.5800

Final joint probability = 0.009280

Normalizing constant:

  - P(F=True | T=True, TB=True) = 2.6150
  - P(TC=True | TB=True) = 0.5930
    Final = 1.550695

P(TB=True | F=True, TC=True) = 0.005984

- Calculating P(TB=False | F=True, TC=True):

Joint probability components:

  - P(TB=False) = 0.9900
  - P(F=True | TB=False, T=True) = 1.0150
  - P(TC=True | TB=False) = 0.0130

Final joint probability = 0.013063

Normalizing constant = 1.550695

P(TB=False | F=True, TC=True) = 0.008424

Final probabilities for TB given F=True, TC=True:

- P(TB=True | F=True, TC=True) = 0.415342
- P(TB=False | F=True, TC=True) = 0.584658

Example 2: Patient has Fever and Bradycardia

Diagnosing Typhoid (T)

- Calculating P(T=True | F=True, B=True):

Joint probability components:

  - P(T=True) = 0.0010
  - P(B=True | T=True) = 0.4000
  - P(F=True | T=True, TB=True) = 2.0000

Final joint probability = 0.000800

Normalizing constant:

  - P(B=True | T=True) = 0.4005
  - P(F=True | T=True, TB=True) = 2.6150
    Final = 1.047308

P(T=True | F=True, B=True) = 0.000764

- Calculating P(T=False | F=True, B=True):

Joint probability components:

  - P(T=False) = 0.9990
  - P(B=True | T=False) = 0.0005
  - P(F=True | T=False, TB=True) = 0.6150

Final joint probability = 0.000307

Normalizing constant = 1.047308

P(T=False | F=True, B=True) = 0.000293

Final probabilities for T given F=True, B=True:

- P(T=True | F=True, B=True) = 0.722548
- P(T=False | F=True, B=True) = 0.277452

Diagnosing Tuberculosis (TB)

- Calculating P(TB=True | F=True, B=True):

Joint probability components:

  - P(TB=True) = 0.0100
  - P(B=True | T=True) = 0.4005
  - P(F=True | TB=True, T=True) = 1.6000

Final joint probability = 0.006408

Normalizing constant:

  - P(B=True | T=True) = 0.4005
  - P(F=True | T=True, TB=True) = 2.6150
    Final = 1.047308

P(TB=True | F=True, B=True) = 0.006119

- Calculating P(TB=False | F=True, B=True):

Joint probability components:

  - P(TB=False) = 0.9900
  - P(B=True | T=True) = 0.4005
  - P(F=True | TB=False, T=True) = 1.0150

Final joint probability = 0.402442

Normalizing constant = 1.047308

P(TB=False | F=True, B=True) = 0.384264

Final probabilities for TB given F=True, B=True:

- P(TB=True | F=True, B=True) = 0.015673
- P(TB=False | F=True, B=True) = 0.984327

6. Discussion: Correlation between Fever and Tachycardia in TB Patients

This observation does not call into question the validity of the results, but it does highlight a limitation in the structure of the Bayesian network used in the previous section.

In the network, Fever (F) and Tachycardia (TC) are conditionally independent given Tuberculosis (TB). That is, TB influences both F and TC separately, but there is no direct dependency modeled between F and TC. This means that once we know whether TB is present, the presence or absence of Fever gives no additional information about Tachycardia, and vice versa.

However, the clinical observation that TB tends to cause fever and tachycardia together implies that there is a dependency between these two symptoms when TB is present. This suggests that the assumption of conditional independence does not hold perfectly in reality.

Therefore, while the computed probabilities in the previous section are correct under the assumptions of the model, the model may not accurately capture all real-world interactions. In particular, the co-occurrence of fever and tachycardia in TB patients could be underestimated or misrepresented.

To better reflect this association, the model could be improved by adding a direct connection between Fever and Tachycardia, or by introducing an intermediate latent variable (e.g., “Systemic Infection Severity”) that influences both symptoms simultaneously.

Exercise 1.5.

Imagine that we are in the 1960s or 1970s, when Bayesian networks had not yet been invented and the only probabilistic diagnostic method was the naïve Bayesian one. Solve the previous exercise, pointing out for each section the **differences** between this model and the Bayesian network.

Solution 1.5.

1. Model Structure

In Naive Bayes, we assume all symptoms are conditionally independent given the diseases. The structure is simpler than a Bayesian network - we only need:

- Prior probabilities for each disease
- Conditional probabilities for each symptom given each disease
- Baseline probabilities for symptoms when no disease is present

2. Variable Values:

All variables are binary (True/False):

- T: Typhoid
- TB: Tuberculosis
- F: Fever
- B: Bradycardia
- TC: Tachycardia

3. Conditional Probabilities:

- **Prior probabilities:**
  - P(T=True) = 0.0010
  - P(TB=True) = 0.0100
- **Conditional probabilities P(Symptom=True | Disease=True):**
  - F (Fever/Bradycardia/Tachycardia):
    - P(F=True | T=True) = 1.0000
    - P(F=True | TB=True) = 0.6000
  - B (Bradycardia):
    - P(B=True | T=True) = 0.4000
    - P(B=True | TB=True) = 0.0000
  - TC (Tachycardia):
    - P(TC=True | T=True) = 0.0000
    - P(TC=True | TB=True) = 0.5800
- **Baseline probabilities (no disease):**
  - P(F=True) = 0.0150

- P(B=True) = 0.0005
- P(TC=True) = 0.0130

4. Hypotheses:

1. Diseases are independent (same as Bayesian network)
2. Symptoms are conditionally independent given the diseases (Naive Bayes assumption)
3. Each symptom's probability depends only on the presence/absence of diseases
4. Baseline rates apply when no disease is present

These hypotheses are different from the Bayesian network because:

- We don’t model causal relationships between symptoms.
- We don’t use an OR gate for fever – each disease contributes independently.
- The model is simpler but less accurate in representing real-world relationships.

5. Diagnosis Table (showing two detailed calculations)

Example 1: Patient has Fever and Tachycardia

Performing diagnosis using Naive Bayes:

- Calculating P(T=True | {'F': True, 'TC': True}):
    - P(F=True | T=True) = 1.0000
    - P(TC=True | T=True) = 0.0000
    - P(F=True | T=False) = 0.0150
    - P(TC=True | T=False) = 0.0130

P(T=True | {'F': True, 'TC': True}) = 0.000000

P(T=True | {'F': True, 'TC': True}) = 0.000000

P(T=False | {'F': True, 'TC': True}) = 1.000000

- Calculating P(TB=True | {'F': True, 'TC': True}):
    - P(F=True | TB=True) = 0.6000
    - P(TC=True | TB=True) = 0.5800
    - P(F=True | TB=False) = 0.0150
    - P(TC=True | TB=False) = 0.0130

P(TB=True | {'F': True, 'TC': True}) = 0.947441

P(TB=True | {'F': True, 'TC': True}) = 0.947441

P(TB=False | {'F': True, 'TC': True}) = 0.052559

Example 2: Patient has Fever and Bradycardia

Performing diagnosis using Naive Bayes:

- Calculating P(T=True | {'F': True, 'B': True}):
    - P(F=True | T=True) = 1.0000
    - P(B=True | T=True) = 0.4000
    - P(F=True | T=False) = 0.0150
    - P(B=True | T=False) = 0.0005

P(T=True | {'F': True, 'B': True}) = 0.981613

P(T=True | {'F': True, 'B': True}) = 0.981613

P(T=False | {'F': True, 'B': True}) = 0.018387

- Calculating P(TB=True | {'F': True, 'B': True}):
    - P(F=True | TB=True) = 0.6000
    - P(B=True | TB=True) = 0.0000
    - P(F=True | TB=False) = 0.0150
    - P(B=True | TB=False) = 0.0005

P(TB=True | {'F': True, 'B': True}) = 0.000000

P(TB=True | {'F': True, 'B': True}) = 0.000000

P(TB=False | {'F': True, 'B': True}) = 1.000000

6. Association between Fever and Tachycardia in TB

In the Naive Bayes model, Fever and Tachycardia are conditionally independent given TB, just like in the Bayesian network. However, the way we calculate probabilities is different:

1. In Naive Bayes, we multiply the individual probabilities.
2. In the Bayesian network, we use the actual causal relationships.

The Naive Bayes model is simpler but less accurate because it doesn’t capture the physiological connection between these symptoms in TB patients.

References

[1] F. J. Díez. *Introducción a los modelos gráficos probabilistas*. UNED, Madrid, 2007.