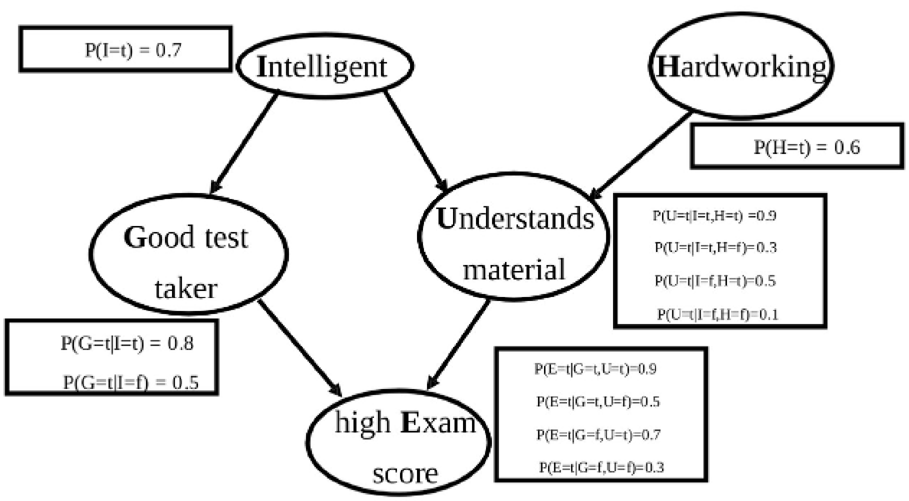
Knowledge Representation and Reasoning
Homework 1: Bayesian Networks (BNs)

Cătălin-Alexandru Rîpanu IA1-A
Automatic Control and Computers
National University of Science and Technology POLITEHNICA Bucharest
E-mail:`catalin.ripanu@upb.ro`
Laboratory Assistant: Mihai Trăscău

15 December 2024

1 Problem I

An instructor wants to determine whether a student has understood the material, based on the exam score, as the Bayes net in figure bellow. Whether the student scores high on the exam is influenced both by whether he/she is a good test taker, and whether he/she understood the material. Both of those, in turn, are influenced by whether he/she is intelligent; whether he/she understood the material is also influenced by whether the student is a hard worker. All the random variables are Boolean.



- a) Manually derive the probability that a student who did well on the test is intelligent, that is, compute the probability P (I=true | E=true). It is mandatory to express the probability algebraically, before filling in the numbers. We start by using the definition of conditional probability:

$$P(I = \text{true} \mid E = \text{true}) = \frac{P(I = \text{true}, E = \text{true})}{P(E = \text{true})}$$

Numerator: $P(I = \text{true}, E = \text{true})$
This is the joint probability of I being true and E being true. We expand this joint probability by considering all possible configurations of G, H, and U:

$$P(I = \text{true}, E = \text{true}) = \sum_{G,H,U} P(I = \text{true}) \cdot P(G \mid I = \text{true}) \cdot P(H) \cdot P(U \mid I = \text{true}, H) \cdot P(E = \text{true} \mid G, U)$$

Denominator: $P(E = \text{true})$
This is the total probability of E being true, regardless of I:

$$P(E = \text{true}) = \sum_{I,G,H,U} P(I) \cdot P(G \mid I) \cdot P(H) \cdot P(U \mid I, H) \cdot P(E = \text{true} \mid G, U)$$

Thus, the final expression for the conditional probability is:

$$P(I = \text{true} \mid E = \text{true}) = \frac{\sum_{G,H,U} P(I = \text{true}) \cdot P(G \mid I = \text{true}) \cdot P(H) \cdot P(U \mid I = \text{true}, H) \cdot P(E = \text{true} \mid G, U)}{\sum_{I,G,H,U} P(I) \cdot P(G \mid I) \cdot P(H) \cdot P(U \mid I, H) \cdot P(E = \text{true} \mid G, U)}$$

Step-by-Step Calculation

Given Probabilities:

$$P(I = \text{true}) = 0.7, \quad P(H = \text{true}) = 0.6$$

$$P(G = \text{true} \mid I = \text{true}) = 0.8, \quad P(G = \text{true} \mid I = \text{false}) = 0.5$$

$$P(U = \text{true} \mid I = \text{true}, H = \text{true}) = 0.9, \quad P(U = \text{true} \mid I = \text{true}, H = \text{false}) = 0.3$$

$$P(U = \text{true} \mid I = \text{false}, H = \text{true}) = 0.5, \quad P(U = \text{true} \mid I = \text{false}, H = \text{false}) = 0.1$$

$$P(E = \text{true} \mid G = \text{true}, U = \text{true}) = 0.9, \quad P(E = \text{true} \mid G = \text{true}, U = \text{false}) = 0.5$$

$$P(E = \text{true} \mid G = \text{false}, U = \text{true}) = 0.7, \quad P(E = \text{true} \mid G = \text{false}, U = \text{false}) = 0.3$$

1. Compute $P(U = \textbf{true} \mid I = \textbf{true})$:
$P(U = \text{true} \mid I = \text{true}) = P(U = \text{true} \mid I = \text{true}, H = \text{true}) \cdot P(H = \text{true}) + P(U = \text{true} \mid I = \text{true}, H = \text{false}) \cdot P(H = \text{false})$

$$P(U = \text{true} \mid I = \text{true}) = 0.9 \cdot 0.6 + 0.3 \cdot 0.4 = 0.54 + 0.12 = 0.66$$

2. Compute $P(E = \textbf{true} \mid I = \textbf{true})$:

$$P(E = \text{true} \mid I = \text{true}) = \sum_{G,U} P(E = \text{true} \mid G, U) \cdot P(G \mid I = \text{true}) \cdot P(U \mid I = \text{true})$$

$$P(E = \text{true} \mid I = \text{true}) = 0.9 \cdot 0.8 \cdot 0.66 + 0.5 \cdot 0.8 \cdot 0.34 + 0.7 \cdot 0.2 \cdot 0.66 + 0.3 \cdot 0.2 \cdot 0.34$$

$$P(E = \text{true} \mid I = \text{true}) = 0.4752 + 0.136 + 0.0924 + 0.0204 = 0.724$$

3. Compute $P(U = \textbf{true} \mid I = \textbf{false})$:

$$P(U = \text{true} \mid I = \text{false}) = P(U = \text{true} \mid I = \text{false}, H = \text{true}) \cdot P(H = \text{true}) + P(U = \text{true} \mid I = \text{false}, H = \text{false}) \cdot P(H = \text{false})$$

$$P(U = \text{true} \mid I = \text{false}) = 0.5 \cdot 0.6 + 0.1 \cdot 0.4 = 0.3 + 0.04 = 0.34$$

4. Compute $P(E = \textbf{true} \mid I = \textbf{false})$:

$$P(E = \text{true} \mid I = \text{false}) = \sum_{G,U} P(E = \text{true} \mid G, U) \cdot P(G \mid I = \text{false}) \cdot P(U \mid I = \text{false})$$

$$P(E = \text{true} \mid I = \text{false}) = 0.9 \cdot 0.5 \cdot 0.34 + 0.5 \cdot 0.5 \cdot 0.66 + 0.7 \cdot 0.5 \cdot 0.34 + 0.3 \cdot 0.5 \cdot 0.66$$

$$P(E = \text{true} \mid I = \text{false}) = 0.153 + 0.165 + 0.119 + 0.099 = 0.536$$

5. Compute $P(E = \textbf{true})$:

$$P(E = \text{true}) = P(E = \text{true} \mid I = \text{true}) \cdot P(I = \text{true}) + P(E = \text{true} \mid I = \text{false}) \cdot P(I = \text{false})$$

$$P(E = \text{true}) = 0.724 \cdot 0.7 + 0.536 \cdot 0.3 = 0.5068 + 0.1608 = 0.6676$$

6. Final Calculation:

$$P(I = \text{true} \mid E = \text{true}) = \frac{P(E = \text{true} \mid I = \text{true}) \cdot P(I = \text{true})}{P(E = \text{true})}$$

$$P(I = \text{true} \mid E = \text{true}) = \frac{0.724 \cdot 0.7}{0.6676} = \frac{0.5068}{0.6676} \approx 0.7591372$$

Conclusion:

The probability that a student is intelligent given that they scored high on the exam is approximately 75.91%.

- b) Verify your manually derived answer against a computed one. Model the given Bayesian Network using the Junction Tree implementation you implemented in the lab sessions. Retrieve the answer for the query and check it against your manual computation.

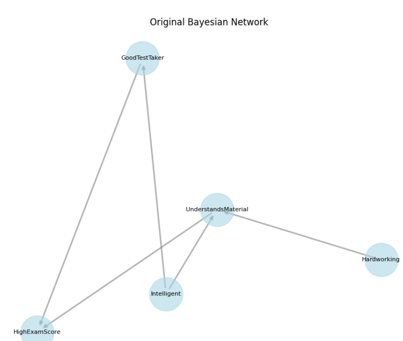


Figure 1: Initial Bayesian Network

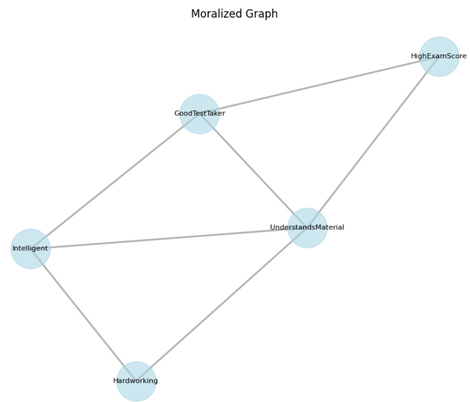


Figure 2: Moralized Bayesian Network

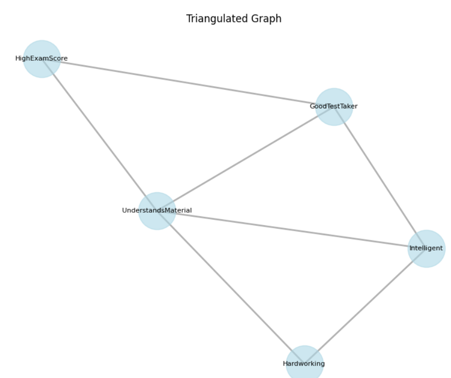


Figure 3: Triangulated Bayesian Network

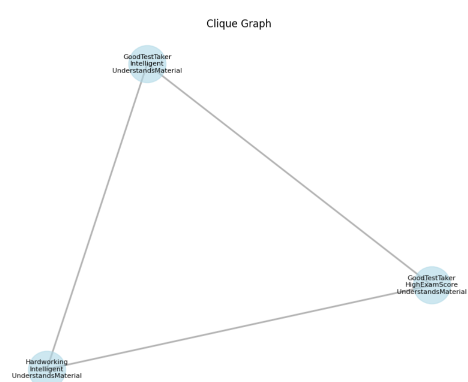


Figure 4: Bayesian Network Clique

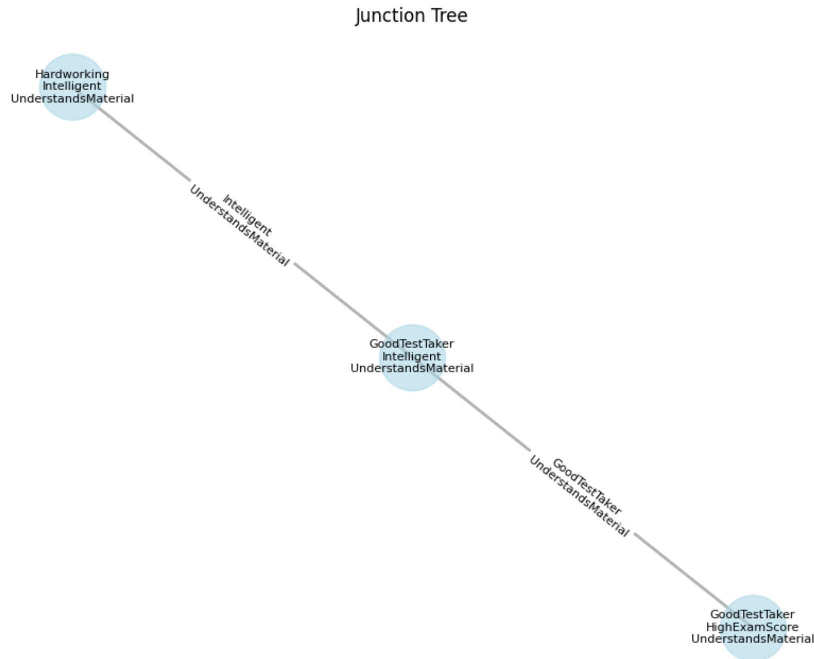


Figure 5: Bayesian Network Junction Tree

Code Output:

```
Moralization complete
Triangulation complete
Elimination order: ['Hardworking', 'Intelligent', 'GoodTestTaker', '
    UnderstandsMaterial', 'HighExamScore']

Found maximal cliques: [['UnderstandsMaterial', 'GoodTestTaker', '
    HighExamScore'], ['UnderstandsMaterial', 'GoodTestTaker', '
    Intelligent'], ['UnderstandsMaterial', 'Hardworking', '
    Intelligent']]

Junction Tree Structure:
Nodes (Cliques): {0: ['GoodTestTaker', 'HighExamScore', '
    UnderstandsMaterial'], 1: ['GoodTestTaker', 'Intelligent', '
    UnderstandsMaterial'], 2: ['Hardworking', 'Intelligent', '
    UnderstandsMaterial']}
Edges (Separators): {(0, 1): ['GoodTestTaker', 'UnderstandsMaterial
    '], (1, 2): ['Intelligent', 'UnderstandsMaterial']}

Query {'query': {'Intelligent': '1'}, 'cond': {'HighExamScore':
    '1'}, 'answer': 0.7591372} OK. Answer is 0.759137, given result
    is 0.759137
```

- c) For the given Bayesian network, determine whether the following statements about conditional independence are **true** or **false**. Provide explanations for each answer.
  1. **G and U are independent. FALSE**
     Explanation: [I influence both, introducing dependency.]
  2. **G and U are conditionally independent given I, E, and H. FALSE**
     Explanation: [Although I is observed, there is E which introduces an active path between G and U (common effect). H is useless in this scenario]
  3. **G and U are conditionally independent given I and H. TRUE**
     Explanation: [I is observed, so G <- I -> U is blocked. Because E is not observed, G -> E <- U is also blocked.]
  4. **E and H are conditionally independent given U. TRUE**
     Explanation: [There are 2 paths from H to E containing U, so if U is observed, it is enough to say both are blocked, meaning H and E are independent.]
  5. **E and H are conditionally independent given U, I, and G. TRUE**
     Explanation: [As both paths from E and H contain U, and U is observed, it is sufficient to conclude the independence between H and E. Both I and G are useless here.]

2 Problem 2

A patient visits a doctor with a medical condition. The doctor considers three possible diseases as potential causes: D_1, D_2, and D_3. These diseases are marginally independent of each other. The presence of four symptoms S_1, S_2, S_3, and S_4 are related to the diseases as follows:

- S_1 depends only on D_1.
- S_2 depends on both D_1 and D_2.
- S_3 depends on both D_1 and D_3.
- S_4 depends only on D_3.

All variables (diseases and symptoms) are Boolean.

- a) Draw the Bayesian network for this problem.

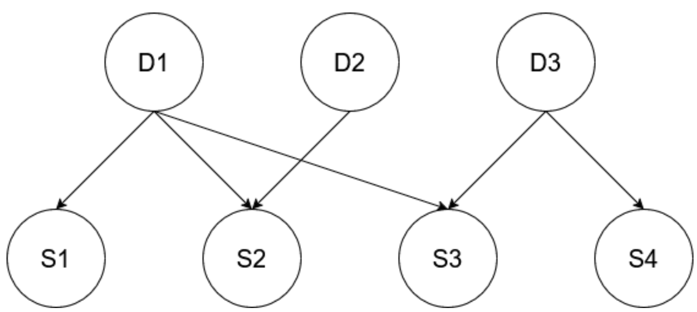


Figure 6: Bayesian Network Structure

- b) What is the number of independent parameters that is required to describe this joint distribution? Explain your answer.

  Diseases:

  Each disease (D_1, D_2, D_3) has one parameter representing $P(D = \text{true})$:

$$\text{Number of parameters for diseases: } 1 + 1 + 1 = 3$$

Symptoms:

- S_1: Depends on D_1 (1 parent). For 1 parent, there are 2 possible values of D_1 ($D_1 = \text{true}$ and $D_1 = \text{false}$). Hence, S_1 requires 2 parameters.
- S_2: Depends on D_1 and D_2 (2 parents). For 2 parents, there are $2^2 = 4$ possible combinations of values (D_1, D_2). Hence, S_2 requires 4 parameters.
- S_3: Depends on D_1 and D_3 (2 parents). For 2 parents, there are $2^2 = 4$ possible combinations of values (D_1, D_3). Hence, S_3 requires 4 parameters.
- S_4: Depends on D_3 (1 parent). For 1 parent, there are 2 possible values of D_3 ($D_3 = \text{true}$ and $D_3 = \text{false}$). Hence, S_4 requires 2 parameters.

Total Parameters:

Adding the parameters for diseases and symptoms:

$$\text{Total Parameters} = 1 + 1 + 1 + 2 + 4 + 4 + 2 = 15$$

Conclusion:

The Bayesian network for this problem requires a total of 15 parameters.

- c) Assume there is no conditional independence between the variables, how many independent parameters would be required then? Explain your answer.

  Total Variables:

  The Bayesian network contains:

  $$3 \text{ diseases} + 4 \text{ symptoms} = 7 \text{ Boolean variables.}$$

  Joint Distribution Without Independence:

  Without assuming independence, the joint distribution would require:

  $$2^7 - 1 = 127 \text{ parameters.}$$

  This is because a full joint probability distribution for 7 Boolean variables has $2^7 = 128$ entries, and the sum of all probabilities must equal 1, leaving $128 - 1 = 127$ free parameters.

- d) What is the Markov Blanket of variable S2?

  Markov Blanket for S_2:

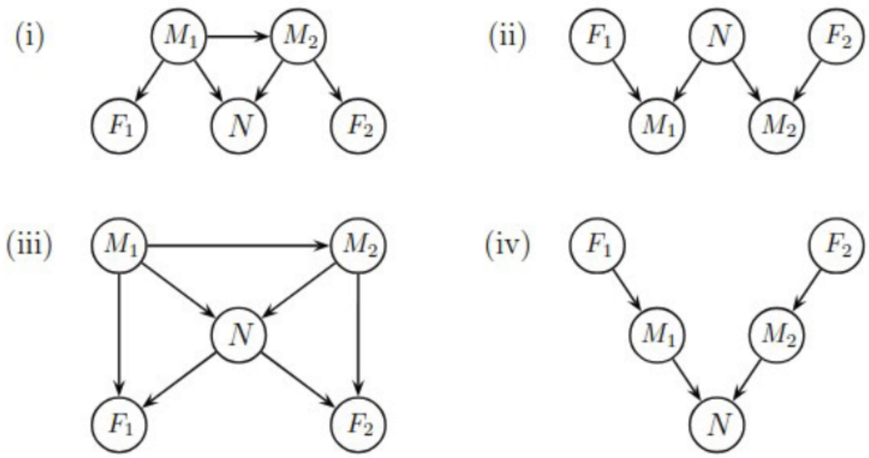
  - **Parents:** D_1, D_2
  - **Children:** None
  - **Parents of Children:** N/A

  Thus, the Markov Blanket of S_2 is $\{D_1, D_2\}$.

3 Problem 3

Two astronomers in two different parts of the world, make measurements M1 and M2 of the number of stars N in some small regions of the sky, using their telescopes. Normally, there is a small possibility of error by up to one star in each direction. Each telescope can be, with a much smaller probability, badly out of focus (events F1 and F2). In such a case the scientist will undercount by three or more stars or, if N is less than three, fail to detect any stars at all.

Consider the four Bayesian Networks shown below:



- a) Which of them correctly, but not necessarily efficiently, represents the above information? (Multiple answers may be possible). Justify your answer
  - D-separation states that, in the Bayesian network structure, F_1 and N are conditionally independent given M_1. However, if M_1 is known to be 1000 in the described situation, then learning that N is 2000 increases the probability that F_1 is true to one. Thus, F_1 and N are not necessarily conditionally independent given M_1.
  - D-separation states that, in the Bayesian network structure, M_1 and M_2 are independent if N is not given. However, if F_1 and F_2 are known to be false in the described situation, then learning that M_1 is 1000 increases the probability that N is in the range 999-1001 to one, which in turn increases the probability that M_2 is in the range 998-1002 to one. Thus, M_1 and M_2 are not necessarily independent if N is not given.

    Thus, (ii) and (iii) are the correct networks.

- b) Which is the best network? Justify your answer.
  Network (ii) is the best:
  - It directly represents the dependencies without introducing unnecessary connections.
  - It efficiently captures the relationships with minimal complexity.