Inteligência Artificial - Lista de Exercícios de Redes Bayesianas

Roberto Sérgio Ribeiro de Meneses - 520403

30/11/2024

1. (Monty Hall Problem). Em um programa de televisão o participante é convidado a escolher entre três portas e atrás de uma delas encontra-se um prêmio (por exemplo, um carro). Após a escolha do participate, o apresentador abre uma das portas restantes e esta, obviamente, não contem o prêmio. A seguir, o apresentador oferece ao participante a opção de trocar a porta que escolheu pela outra porta restante. Qual deve ser a escolha do participante para que a sua chance de ganhar seja maximizada ?

Resposta

O participante deve trocar de porta para maximizar sua chance de ganhar.

Explicação

Inicialmente, o participante tem uma probabilidade de $\frac{1}{3}$ de escolher a porta correta e uma probabilidade de $\frac{2}{3}$ de escolher uma porta errada.

Após a escolha inicial, o apresentador abre uma das portas restantes que não contém o prêmio. Se a escolha inicial do participante foi errada (o que ocorre com probabilidade $\frac{2}{3}$), a troca levará à porta correta. Se a escolha inicial foi correta (probabilidade $\frac{1}{3}$), a troca levará à porta errada.

Dessa forma, se o participante mantiver sua escolha inicial, sua chance de vitória permanece $\frac{1}{3}$. Se ele optar por trocar, sua chance de ganhar aumenta para $\frac{2}{3}$.

2. Considere a seguinte rede Bayesiana

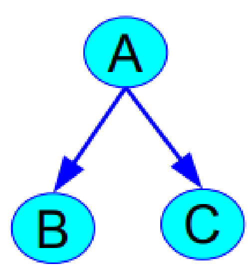

Figura 1: Rede Bayesiana do problema.

Onde:

$$P(A) = 0.5, \quad P(B|A) = P(B|\neg A) = 0.2, \quad P(C|A) = 0.8, \quad P(C|\neg A) = 0.4$$

Calcule:

a) $P(B)$

b) $P(B|C)$

c) $P(C|B)$

Resolução

a) Cálculo de $P(B)$

A probabilidade total de B é obtida pela Regra da Probabilidade Total:

$$P(B) = P(B|A)P(A) + P(B|\neg A)P(\neg A)$$

Substituindo os valores:

$$P(B) = (0.2 \times 0.5) + (0.2 \times 0.5)$$

$$P(B) = 0.1 + 0.1 = 0.2$$

b) Cálculo de $P(B|C)$

Usamos a definição de probabilidade condicional:

$$P(B|C) = \frac{P(B \cap C)}{P(C)}$$

Calculamos $P(B \cap C)$:

$$P(B \cap C) = P(B|A)P(C|A)P(A) + P(B|\neg A)P(C|\neg A)P(\neg A)$$

$$P(B \cap C) = (0.2 \times 0.8 \times 0.5) + (0.2 \times 0.4 \times 0.5)$$

$$P(B \cap C) = (0.08) + (0.04) = 0.12$$

Agora, calculamos $P(C)$:

$$P(C) = P(C|A)P(A) + P(C|\neg A)P(\neg A)$$

$$P(C) = (0.8 \times 0.5) + (0.4 \times 0.5)$$

$$P(C) = 0.4 + 0.2 = 0.6$$

Agora, encontramos $P(B|C)$:

$$P(B|C) = \frac{0.12}{0.6} = 0.2$$

c) Cálculo de $P(C|B)$

Usamos novamente a definição de probabilidade condicional:

$$P(C|B) = \frac{P(B \cap C)}{P(B)}$$

Já calculamos $P(B \cap C) = 0.12$ e $P(B) = 0.2$, então:

$$P(C|B) = \frac{0.12}{0.2} = 0.6$$

3. Considere a seguinte rede Bayesiana

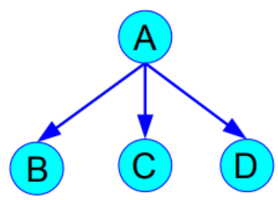

Figura 2: Rede Bayesiana do problema.

Com as seguintes probabilidades:

$$P(A) = 0.5, \quad P(B|A) = P(C|A) = P(D|A) = 0.2$$

$$P(B|\neg A) = P(C|\neg A) = P(D|\neg A) = 0.6$$

Calcule:
a) $P(C|B, A)$
b) $P(A|B, C, D)$
c) $P(C|B)$

Resolução

a) Cálculo de $P(C|B, A)$

Sabemos que, na rede Bayesiana dada, B, C e D são condicionalmente independentes dado A. Assim, podemos escrever:

$$P(C|B, A) = P(C|A)$$

Substituindo o valor fornecido:

$$P(C|B, A) = P(C|A) = 0.2$$

b) Cálculo de $P(A|B, C, D)$

Usamos o Teorema de Bayes:

$$P(A|B, C, D) = \frac{P(B, C, D|A)P(A)}{P(B, C, D)}$$

Primeiro, calculamos $P(B, C, D|A)$:

$$P(B, C, D|A) = P(B|A)P(C|A)P(D|A)$$

$$= (0.2 \times 0.2 \times 0.2) = 0.008$$

Agora, calculamos $P(B, C, D|\neg A)$:

$$P(B, C, D|\neg A) = P(B|\neg A)P(C|\neg A)P(D|\neg A)$$

$$= (0.6 \times 0.6 \times 0.6) = 0.216$$

Usamos a Regra da Probabilidade Total para encontrar $P(B, C, D)$:

$$P(B, C, D) = P(B, C, D|A)P(A) + P(B, C, D|\neg A)P(\neg A)$$

$$P(B, C, D) = (0.008 \times 0.5) + (0.216 \times 0.5)$$

$$= 0.004 + 0.108 = 0.112$$

Agora, encontramos $P(A|B, C, D)$:

$$P(A|B, C, D) = \frac{0.008 \times 0.5}{0.112}$$

$$= \frac{0.004}{0.112} = 0.0357$$

c) Cálculo de $P(C|B)$

Usamos a Regra da Probabilidade Total:

$$P(C|B) = \frac{P(B \cap C)}{P(B)}$$

$$P(C|B) = P(C|A) * P(B|A) * P(A) + P(C|\neg A) * P(B|\neg A) * P(\neg A)$$

Sabemos que:

$$P(B \cap C) = P(C|A) * P(B|A) * P(A) + P(C|\neg A) * P(B|\neg A) * P(\neg A)$$

Calculando: $P(B \cap C) = 0.2$

Ajustando os cálculos:

$$P(B) = P(B|A) * P(A) + P(B|\neg A) * P(\neg A)$$

$$= 0.4$$

Calculando P(C|B)

$$= 0.2/0.4 = 0.5$$

4.Considere a seguinte rede Bayesiana:

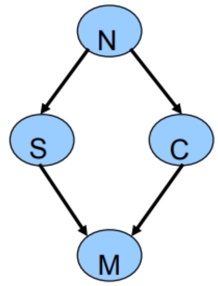

Figura 3: Rede Bayesiana do problema.

Verifique se (selecione entre V ou F):

1. **S $\perp$ N:**
   A resposta é **Falsa (F)**, pois N tem uma aresta direta para S ($N \rightarrow S$).
2. **S $\perp$ C:**
   A resposta é **Falsa (F)**, pois N é um nó comum que conecta tanto S quanto C, criando uma dependência entre as duas variáveis.
3. **S $\perp$ C | N:**
   A resposta é **Verdadeira (V)**, pois ao condicionar N, qualquer dependência entre S e C não tem mais, tornando-as independentes.
4. **S $\perp$ C | M:**
   A resposta é **Falsa (F)**, pois ao condicionar em M , cria-se um caminho ativo de $S->M$ e $C->M$.
5. **S $\perp$ C | M, N:**
   A resposta é **Falsa (F)**, pois ao condicionar em M , cria-se um caminho ativo de $S->M$ e $C->M$.

5.Determine:

1. O número de parâmetros necessários para especificar a distribuição conjunta de 5 variáveis binárias.
2. Assumindo que a relação entre as variáveis é dada pela rede Bayesiana abaixo, determine o número de parâmetros necessários para especificar a distribuição conjunta das 5 variáveis.

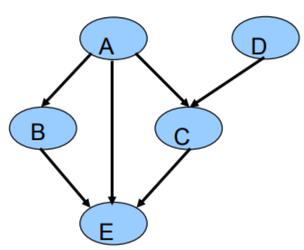

Figura 4: Rede Bayesiana do problema.

Solução

(a) Distribuição Conjunta sem Suposições de Independência

Se temos 5 variáveis binárias X_1, X_2, X_3, X_4, X_5, a distribuição conjunta completa pode ser especificada por uma tabela de probabilidades com $2^5 = 32$ entradas. No entanto, como a soma das probabilidades deve ser 1, precisamos apenas de $32 - 1 = 31$ parâmetros independentes.

(b) Com a Rede Bayesiana

A rede Bayesiana define independências condicionais que reduzem o número de parâmetros necessários. Para cada variável X_i com pais $Pa(X_i)$, precisamos especificar $P(X_i|Pa(X_i))$. Como as variáveis são binárias, cada conjunto de pais gera $2^{\text{número de pais}}$ valores possíveis.

A estrutura dada é:

- A (sem pais): precisa de 1 parâmetro $P(A = 1)$.
- $B|A$: precisa de 2 parâmetros, $(2 - 1) * 2$
- $C|A, D$: precisa de 4 parâmetros, $(2 - 1) * (2 * 2)$.
- D (sem pais): precisa de 1 parâmetro $P(D = 1)$.
- $E|A, B, C$: precisa de 8 parâmetros, $(2 - 1) * (2 * 2 * 2)$.

Somando todos os parâmetros necessários:

$$1 + 2 + 4 + 1 + 8 = 16 \tag{1}$$

Portanto, com a estrutura da rede Bayesiana dada, precisamos de **16 parâmetros** para especificar a distribuição conjunta das 5 variáveis.

6.Enunciado

Considere um conjunto de variáveis utilizadas para modelar um problema utilizando uma rede Bayesiana. A rede Bayesiana deve modelar uma situação onde um motorista tem a opção de voltar para casa após uma festa onde este pode ter ingerido bebida alcoólica. As variáveis em questão são as seguintes:

- **Bêbado** – Variável binária (bêbado ou não)
- **Chovendo** – Variável binária (Está chovendo ou não)
- **Preso** – Variável binária (O motorista será preso ou não)
- **Falha nos freios** – Variável binária (irá ocorrer falha nos freios)
- **Acidente** – Variável binária (O motorista provocará um acidente)
- **Gravidade do acidente** – Três níveis (leve, moderado e grave)

Apresente uma proposta de estrutura para uma rede Bayesiana e determine o número de parâmetros necessários para especificar a probabilidade conjunta de todas as variáveis.

Rede Bayesiana construida

A estrutura da rede Bayesiana pode ser modelada conforme o seguinte grafo:

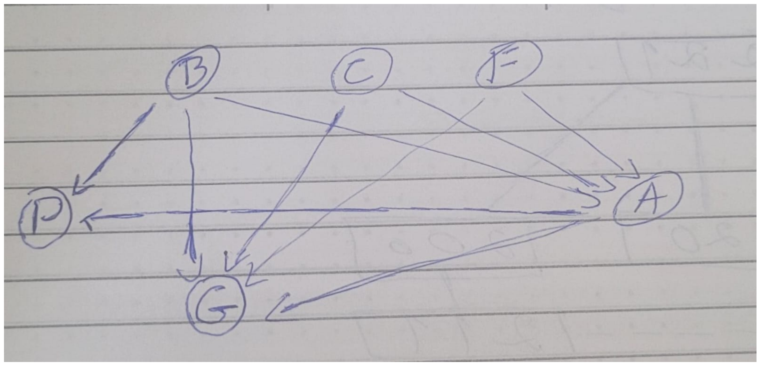

Figura 5: Rede Bayesiana construída.

Agora, calculamos o número de parâmetros para cada variável, considerando suas dependências.

1.Bêbado (B)

- Variável binária (bêbado ou não).
- Não tem pais.
- Número de parâmetros: **1** (probabilidade de estar bêbado, $P(B)$).

2. Chovendo (C)

- Variável binária (está chovendo ou não).
- Não tem pais.
- Número de parâmetros: **1** (probabilidade de estar chovendo, $P(C)$).

3. Falha nos freios (F)

- Variável binária (falha ou não).
- Não tem pais.
- Número de parâmetros: **1** (probabilidade de falha nos freios, $P(F)$).

4. Acidente (A)

- Variável binária (acidente ou não).
- Depende de **B**, **C** e **F**.
- Cada combinação de pais ($2 \times 2 \times 2 = 8$ combinações) requer 1 parâmetro.
- Número de parâmetros: **8** (probabilidades condicionais $P(A \mid B, C, F)$).

5. Gravidade do acidente (G)

- Variável com três níveis (leve, moderado, grave).
- Depende de **B**, **C**, **F** e **A**.
- Cada combinação de pais ($2 \times 2 \times 2 \times 2 = 16$ combinações) requer 2 parâmetros.
- Número de parâmetros: $16 \times (3 - 1) = 32$.

6. Preso (P)

- Variável binária (preso ou não).
- Depende de **B** e **A**.
- Cada combinação de pais (2 estados de **B** e 2 estados de **A**) requer 1 parâmetro.
- Número de parâmetros: $2 \times 2 = 4$ (probabilidades condicionais $P(P \mid B, A)$).

Soma Total de Parâmetros

Agora, somamos os parâmetros de todas as variáveis:

Total de parâmetros necessários:

$$1 + 1 + 1 + 8 + 32 + 4 = 45$$

Observações

- A variável **Gravidade do acidente (G)** contribui com o maior número de parâmetros (32), pois depende de 4 pais (**B**, **C**, **F**, **A**), cada um com 2 estados, resultando em 16 combinações. Para cada combinação, são necessários 2 parâmetros (já que a soma das probabilidades dos 3 níveis deve ser 1).
- A rede proposta tem mais dependências diretas para **G** (depende de **B**, **C**, **F** e **A**), o que aumenta significativamente o número de parâmetros em comparação com uma estrutura mais simples.

Conclusão

O número total de parâmetros necessários para especificar a probabilidade conjunta de todas as variáveis na rede proposta é **45**.

7.Considere as seguintes redes Bayesianas desenvolvidas com o objetivo de modelar o problema de diagnóstico de câncer no pulmão. Nas figuras apresentadas as variáveis são:

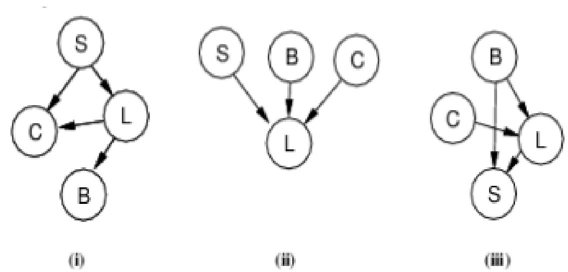

Figura 6: Rede Bayesiana do problema.

- **S** – Fumante (sim ou não)
- **L** – Tem Câncer (sim ou não)
- **B** – Resultado da Biópsia (positivo ou negativo)
- **C** – Tem Tosse (sim ou não)

(a) Das opções listadas, qual apresenta a melhor modelagem do problema em questão?
(b) Qual modelo proposto apresenta o menor número de parâmetros?

Resposta

(a) Melhor Modelagem do Problema

A modelagem deve refletir relações causais realistas. Consideremos:

- **S (Fumante)** influencia o risco de desenvolver **L (Câncer)**.
- **L (Câncer)** é a condição que determina o resultado da **B (Biópsia)** e, além disso, pode causar **C (Tosse)**.

No **Modelo A** além de S influenciar L, temos também uma seta direta de S para C. Essa conexão representa a possibilidade de que o fumo cause tosse por si só (por exemplo, devido à irritação das vias aéreas), sem a intermediação do câncer.

No **Modelo B** Temos uma situação mais simples onde fumar, Biopsia e tosse não tem nenhuma relação entre si e que implicam em cancer, o que em comparação com o modelo A não leve em consideração informações importantes como fumar tb pode causar tosse, além de que nessa rede a biopsia implica em cancer.

Já o **Modelo C** Temos uma situação sendo desenvolvida ao contrário do modelo A, pois nesse modelo a Biopsia que implica em câncer e na tosse e essa ideia não é boa por que é o cancer que leva a biopsia ser positiva ou negativa e não contrário.

Conclusão (a): O **Modelo A** é a melhor modelagem, pois incorpora tanto o efeito do fumo com o câncer tanto quanto o efeito do fumo com a tosse, além de que também é o único modelo que considera que o câncer implica no resultado da biopsia e não o contrário.

(b) Modelo com o Menor Número de Parâmetros

O número de parâmetros necessários para especificar a distribuição condicional de uma variável X com k estados e n pais, onde cada pai i tem p_i estados, é dado por:

$$(k-1) \times \prod_{i=1}^{n} p_i.$$

Vamos calcular para cada modelo:

Para todas as variáveis, considera-se que possuem 2 estados.

Modelo A:

- **S:** sem pais:
$$(2-1) = 1 \quad \text{parâmetro.}$$
- **L:** com pai S (2 combinações):
$$(2-1) \times 2 = 2 \quad \text{parâmetros.}$$
- **B:** com pai L (2 combinações):
$$(2-1) \times 2 = 2 \quad \text{parâmetros.}$$
- **C:** com pais S e L ($2 \times 2 = 4$ combinações):
$$(2-1) \times 4 = 4 \quad \text{parâmetros.}$$

Total para o Modelo A:
$$1 + 2 + 2 + 4 = 9 \quad \text{parâmetros.}$$

Modelo B:

- **S:** sem pais:
$$1 \quad \text{parâmetro.}$$
- **B:** sem pais:
$$1 \quad \text{parâmetro.}$$
- **C:** sem pais:
$$1 \quad \text{parâmetro.}$$
- **L:** com pais S, B e C ($2 \times 2 \times 2 = 8$ combinações):
$$(2-1) \times 8 = 8 \quad \text{parâmetros.}$$

Total para o Modelo B:
$$1 + 1 + 1 + 8 = 11 \quad \text{parâmetros.}$$

Modelo C:

- **S:** com pai L e B (4 combinações):
$$(2-1) \times 4 = 4 \quad \text{parâmetros.}$$
- **L:** com pai C e B (4 combinações):
$$(2-1) \times 4 = 4 \quad \text{parâmetros.}$$

- **B:** com pai L (2 combinações):
$$1 \quad \text{parâmetro.}$$

- **C:** com pai L (2 combinações):
$$1 \quad \text{parâmetro.}$$

Total para o Modelo C:
$$4 + 4 + 1 + 1 = 10 \quad \text{parâmetros.}$$

Conclusão (b): O **Modelo A** apresenta o menor número de parâmetros (9), seguido pelo Modelo C (10 parâmetros) e, por último, o Modelo B (11 parâmetros).

8. Considere a rede Bayesiana a seguir, utilizada para modelar uma situação de eleição.

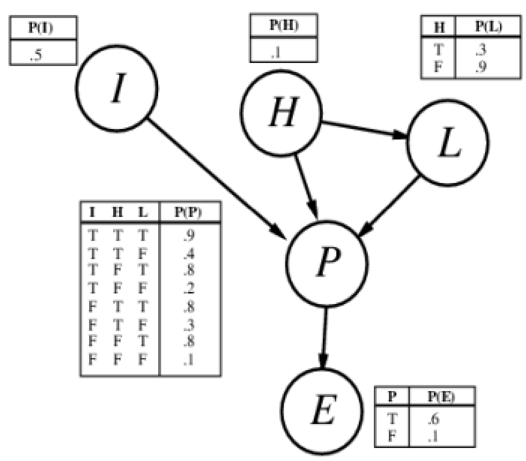

Figura 7: Rede Bayesiana do problema.

As variáveis são:

I – Inteligente, H – Honesto, P – Popular, E – Eleito, L – Muito dinheiro para campanha (Lots of Campaign Funds).

(a) Verifique se as seguintes expressões são verdadeiras, de acordo com a rede apresentada:

$$1.\ P(I, L) \stackrel{?}{=} P(I)\,P(L)$$

$$2.\ P(E \mid P, L) \stackrel{?}{=} P(E \mid P, L, H)$$

$$3.\ P(P \mid I, H) \stackrel{?}{\neq} P(P \mid I, H, L)$$

Estrutura da Rede e Fatorização

A distribuição conjunta fatoriza como:

$$P(I, H, L, P, E) \; = \; P(I)\, P(H)\, P(L)\, P\big(P \mid I, H, L\big)\, P\big(E \mid P\big).$$

Isto significa que:

- I, H e L são variáveis raízes, assumidas independentes entre si.
- P (Popular) depende de I, H e L.
- E (Eleito) depende apenas de P.

1. $P(I, L) = P(I)\, P(L)$ Como I e L são nós iniciais sem qualquer ligação direta ou indireta entre si (a dependência só aparece em P, que é um nó filho comum), eles são marginalmente independentes. Logo,

$$P(I, L) = P(I)\, P(L)$$

é verdadeiro.

2. $P(E \mid P, L) = P(E \mid P, L, H)$ A variável E tem como único pai P. Logo, sabendo P, E torna-se independente de quaisquer outras variáveis, incluindo L e H. Em termos formais:

$$P(E \mid P, L, H) = P(E \mid P).$$

Portanto,

$$P(E \mid P, L) = P(E \mid P, L, H) = P(E \mid P).$$

Assim, a igualdade dada é verdadeira.

3. $P(P \mid I, H) \neq P(P \mid I, H, L)$ A probabilidade de P (Popular) depende dos valores de I, H e também de L. Se não soubermos L, não poderemos calcular a mesma probabilidade que se L fosse conhecido. Em geral,

$$P\big(P \mid I, H, L\big) \; \neq \; P\big(P \mid I, H\big),$$

pois L é pai de P. Assim, esta expressão é verdadeira: a probabilidade não é igual se acrescentamos (ou removemos) L da condição.

(b) Calcule a probabilidade de alguém ser inteligente dado que é honesto, teve pouco dinheiro para campanha ($-L$) e foi eleito.

Queremos calcular:

$$P(I \mid H, -L, E). \tag{2}$$

Regra de Bayes e Soma sobre P Utilizando a regra de Bayes, podemos escrever:

$$P(I \mid H, -L, E) = \frac{P(I, H, -L, E)}{P(H, -L, E)}. \tag{3}$$

Expandindo os termos da distribuição conjunta:

$$P(I, H, -L, E) = \sum_{p \in \{T, F\}} P(I)P(H)P(-L)P(P = p \mid I, H, -L)P(E \mid P = p). \tag{4}$$

Analogamente, o denominador é a soma dos casos para I e $-I$:

$$P(H, -L, E) = \sum_{i' \in \{T, F\}} \sum_{p \in \{T, F\}} P(i')P(H)P(-L)P(P = p \mid i', H, -L)P(E \mid P = p). \tag{5}$$

Exemplo Numérico

Usando as probabilidades fornecidas:

$$\begin{aligned} P(I) &= 0.5, \\ P(H) &= 0.1, \\ P(L) &= 0.3, \\ P(-L) &= 0.7. \end{aligned}$$

E as probabilidades condicionais:

$$\begin{aligned} P(E \mid P) &= 0.6, \\ P(E \mid -P) &= 0.1. \end{aligned}$$

Cálculo do Numerador:

$$\begin{aligned} P(I, H, -L, E) &= P(I)P(H)P(-L) \sum_{p \in \{T,F\}} P(P = p \mid I, H, -L)P(E \mid P = p) \\ &= (0.5)(0.1)(0.7)\big[(0.4)(0.6) + (0.6)(0.1)\big] \\ &= (0.5)(0.1)(0.7)(0.24 + 0.06) \\ &= 0.5 \times 0.1 \times 0.7 \times 0.3 \\ &= 0.0105. \end{aligned}$$

Cálculo do Denominador:

$$\begin{aligned} P(H, -L, E) &= \sum_{i' \in \{T,F\}} P(i')P(H)P(-L) \sum_{p \in \{T,F\}} P(P = p \mid i', H, -L)P(E \mid P = p) \\ &= (0.5)(0.1)(0.7)(0.24 + 0.06) + (0.5)(0.1)(0.7)(0.18 + 0.07) \\ &= 0.0105 + 0.00875 \\ &= 0.01925. \end{aligned}$$

Probabilidade Final:

$$P(I \mid H, L, E) = \frac{0.0105}{0.01925} \approx 0.545. \tag{6}$$

9. enunciado

Três times de futebol, A, B e C, jogam um contra o outro. Cada jogo tem três resultados possíveis (vitória de cada time e empate). Cada time tem um nível fixo de qualidade (número inteiro de 0 a 3) e este nível influencia probabilisticamente no resultado da partida.

(a) Estrutura da Rede Bayesiana

A estrutura da rede Bayesiana que modela o problema é representada pelo seguinte grafo:

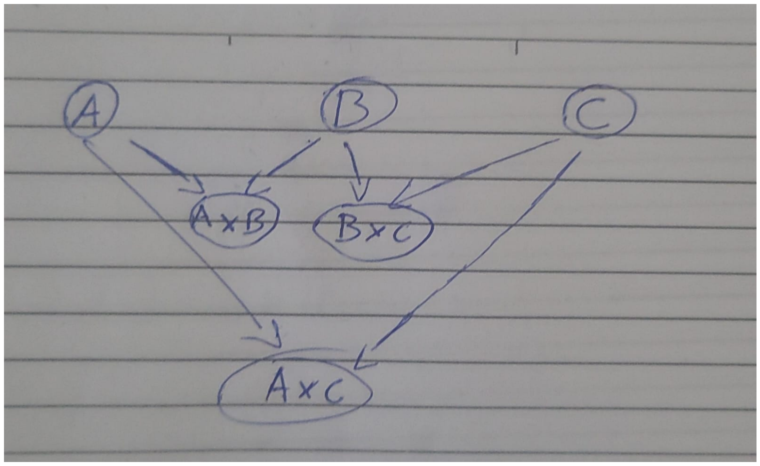

Figura 8: Rede Bayesiana que construi.

Os nós representam os times e os jogos disputados entre eles, considerando a influência da qualidade de cada time sobre o resultado da partida.

(b) Tabelas de Probabilidades Condicionais

Cada time possui uma probabilidade de 25% de ser classificado em um dos 4 níveis possíveis de qualidade, que são representados pelos valores 0, 1, 2 e 3. Essas probabilidades são distribuídas de forma uniforme, ou seja:

$$P(\text{Nível de um time} = 0) = 25\%, \quad P(\text{Nível de um time} = 1) = 25\%, \quad P(\text{Nível de um time} = 2) = 25\%,$$

$$P(\text{Nível de um time} = 3) = 25\%$$

Ou seja, cada time tem uma chance igual de ser alocado em qualquer um dos níveis, independentemente dos outros times.

As probabilidades de cada resultado (Vitória de um time ou Empate) dependem dos níveis dos times. Consideramos as seguintes distribuições de probabilidade:

- Se os níveis dos times forem iguais:

$$P(\text{Vitória do time com nível maior}) = 25\%, \quad P(\text{Empate}) = 50\%, \quad P(\text{Vitória do outro time}) = 25\%$$

- Se os níveis diferirem em 1:

$$P(\text{Vitória do time mais forte}) = 40\%, \quad P(\text{Empate}) = 40\%, \quad P(\text{Vitória do mais fraco}) = 20\%$$

- Se os níveis diferirem em 2:

$$P(\text{Vitória do time mais forte}) = 55\%, \quad P(\text{Empate}) = 30\%, \quad P(\text{Vitória do mais fraco}) = 15\%$$

- Se os níveis diferirem em 3:

$$P(\text{Vitória do time mais forte}) = 70\%, \quad P(\text{Empate}) = 20\%, \quad P(\text{Vitória do mais fraco}) = 10\%$$

As tabelas abaixo mostram as probabilidades condicionais dos resultados para os jogos $A \times B$, $A \times C$ e $B \times C$, dependendo dos níveis de qualidade dos times.

Jogo $A \times B$

Nível de $A vs$ Nível de B	$P(A$ vence$)$	$P($Empate$)$	$P(B$ vence$)$
$0vs0$	0.25	0.50	0.25
$0vs1$	0.20	0.40	0.40
$0vs2$	0.15	0.30	0.55
$0vs3$	0.10	0.20	0.70
$1vs0$	0.40	0.40	0.20
$1vs1$	0.25	0.50	0.25
$1vs2$	0.20	0.40	0.40
$1vs3$	0.15	0.30	0.55
$2vs0$	0.55	0.30	0.15
$2vs1$	0.40	0.40	0.20
$2vs2$	0.25	0.50	0.25
$2vs3$	0.20	0.40	0.40
$3vs0$	0.70	0.20	0.10
$3vs1$	0.55	0.30	0.15
$3vs2$	0.40	0.40	0.20
$3vs3$	0.25	0.50	0.25

Jogo $A \times C$

(Mesmo formato da tabela acima)

Jogo $B \times C$

(Mesmo formato da tabela acima)