



**ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO**  
**Departamento de Engenharia de Computação e Sistemas Digitais**

---

---

**PCS - 2059 Inteligência Artificial**

Prof. Jaime Simão Sichman

**2<sup>a</sup> Lista de Exercícios**

**Planejamento**

1. Discuta as vantagens e desvantagens entre um plano totalmente ordenado e um plano parcialmente ordenado.

Um plano parcialmente ordenado é mais flexível, uma vez que ele não define uma ordem total entre suas ações. Assim, o plano pode se manter aplicável caso, eventualmente, alguma nova restrição que introduza a necessidade de se estabelecer uma ordem entre duas ações que não tinham essa relação (direta ou indiretamente) no plano original. Além disso, teoricamente, este tipo de plano pode ser construído consumindo menos recursos de processamento (mais rápido), pois neste caso tem-se que encontrar e definir menos relações de ordem entre as ações. Entretanto, um plano totalmente ordenado tem como principal vantagem o fato de já estar completo, assim, por exemplo, caso um agente tenha de executá-lo ele não terá que dedicar recursos (processamento) para linearizá-lo. Este tipo de plano também evita a redundância da linearização, por exemplo, dois agentes A e B teriam perder tempo linearizando o mesmo plano se este não já não estiver completamente ordenado.

2. Compare as estratégias progressivas e regressivas para a geração de planos.

Um plano construído de forma progressiva parte do estado inicial do problema para chegar ao objetivo o que funciona muito bem onde o objetivo na verdade é representado por mais de estado (múltiplos objetivos) ou quando o há mais de um estado que possa ser considerado como um objetivo. Neste último caso, durante a construção do plano ao chegar-se em qualquer um dos objetivos é possível dar a

busca como concluída, sem ter que encontrar as demais possíveis soluções. Porém, quando se há um único objetivo a estratégia progressiva têm uma vantagem em problemas onde normalmente há mais links saindo do estado inicial do que do objetivo, assim ela acaba visitando menos estados na busca e consequentemente acaba por encontrar o estado inicial relativamente rápido.

3. Crie um plano não completamente ordenado para o problema (indique possíveis ameaças, caso elas existam).

*Estado Inicial:* NoCarro(pai), NoCarro(filho)

*Objetivo:* Em(pai, trabalho) ^ Em(filho, escola) ^ TanqueCheio(Carro) ^ ¬NoCarro(pai) ^ ¬NoCarro(filho)

*Operadores*

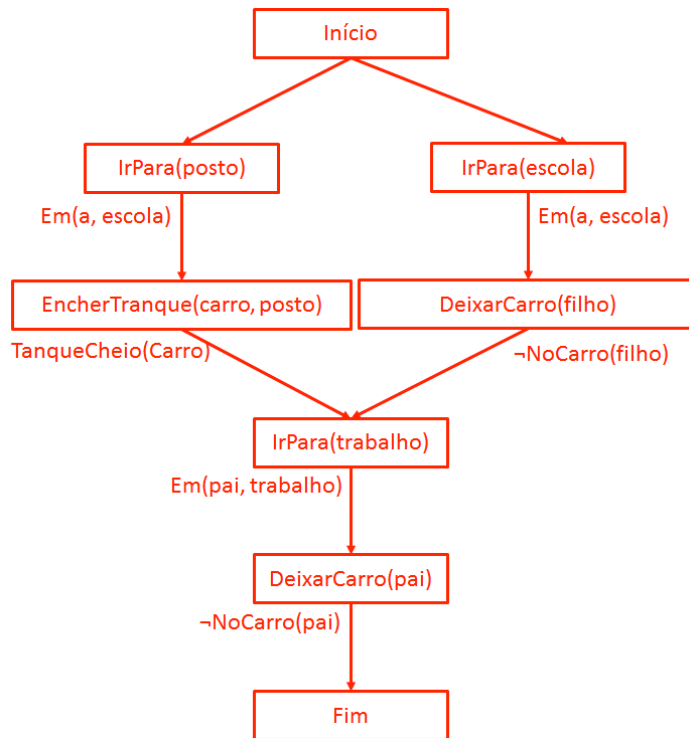
Op(Ação:IrPara(A), Pré-condição: Em(C, B), Efeito: Adicionar:Em(C, A), Remover: Em(C, B))

Op(Ação:DeixarCarro(A), Efeito: Remover: NoCarro(A))

Op(Ação:EncherTanque(Carro, Posto), Pré-condição: Em(A, Posto), Efeito: TanqueCheio(Carro))

**Plano**

- 1) IrPara (posto) ou IrPara(escola)
- 2) EncherTranque(carro, posto) ou DeixarCarro(filho)
- 3) IrPara(trabalho)
- 4) Deixar(pai)



## Sistemas de Produção

1. Quais são os três principais módulos que compõem um sistema de produção e qual a função de cada um deles, resalte também como eles se relacionam.

**Base de regras:** composto por um conjunto de regras no formato se-então que representam o comportamento do sistema, normalmente a base não muda durante a execução do sistema por questões de integridade e coerência do mesmo.

**Memória de trabalho:** composta por fatos que são gerados através da percepção das entidades autônomas dos sistemas (normalmente agentes) e também das regras presentes na base de regras.

**Mecanismo de inferência:** representa implementação da estratégia de unificação e disparo de regras, ou seja, o método de raciocínio utilizado para se gerar novos fatos. Também é responsável pelo encontro e gerenciamento de conflitos de forma pré-determinada (i. e. decisão sobre qual regra deve ser disparada primeiro, quando mais de uma pode ser executada).

**Dinâmica de funcionamento e interação:** Os fatos presentes na memória de trabalho estão em constante atualização através de um ciclo onde novos fatos são gerados através do mecanismo de inferência pelo uso da base de regras. Os novos fatos gerados são usados pelo mecanismo de inferência iniciando um novo ciclo que se perpetua até que não seja possível gerar novos fatos ou que o sistema atinja uma restrição de funcionamento.

2. Dada a base de regras e os fatos a baixo, descreva três rodadas de inferência usando o encadeamento progressivo, em cada rodada a resposta deverá conter: as regras disparáveis, a regra selecionada para a execução e os fatos (incluindo os novos fatos gerados) na base de conhecimento.

Para resolução de conflitos entre duas regras A e B, utilize:

1.  $A > B$  se A for descrita (abaixo) antes de B.
2.  $A > B$  se B foi executada antes de A.

*Base de Regras*

VeículoSustentável: Se motor = elétrico, Então tipo = sustentável

VeículoNãoSustentável: Se motor = diesel, Então tipo = não sustentável

Trem Elétrico: Se andaSobre= Trilhos

E tipo = sustentável

Então veículo = trem elétrico

Ônibus: andaSobre = asfalto

E articulações = 0

Então veículoTamanho = médio

ÔnibusArticulado: andaSobre = asfalto

E articulações > 0

Então veículo = grande

Tróibus: andaSobre = asfalto

E tipo = sustentável

Então veículo = tróibus

Ônibus: andaSobre = asfalto

E tipo = sustentável

Então veículo = ônibus

*Fatos iniciais:*

motor = elétrico, articulações = 2, andaSobre = asfalto

Rodada 1

Regras executáveis:

VeículoSustentável, ÔnibusArticulado

Regra escolhida: VeículoSustentável

Novo fato na memória de trabalho: tipo = sustentável

Rodada 2

Regras executáveis:

VeículoSustentável, ÔnibusArticulado, Trólebus

Regra escolhida: ÔnibusArticulado

Novo fato na memória de trabalho: tamanho = grande

Rodada 3:

Regras executáveis:

VeículoSustentável, ÔnibusArticulado, Trólebus

Regra escolhida: Trólebus

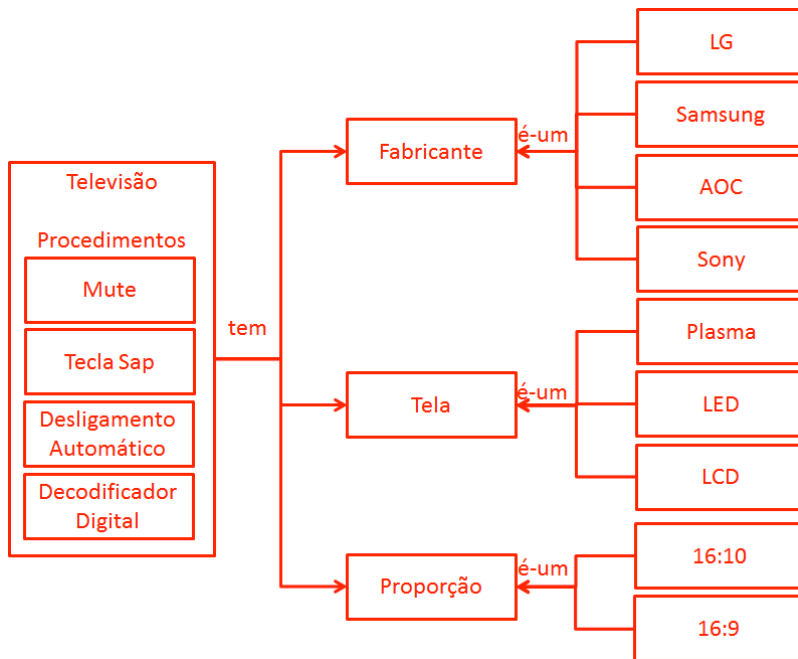
Novo fato na memória de trabalho: veículo = trólebus

Fatos: motor = elétrico, articulações = 2, andaSobre = asfalto, veículo = trólebus, tamanho = grande, tipo = sustentável

### 3. Dadas das seguintes informações

- Existem 4 tipos de televisão Plasma, LED, LCD e CRT;
- Existem 3 tipos de proporção de telas: 16:10, 16:9, e 4:3;
- Existem 4 fabricantes de televisão: LG, AOC, Samsung e Sony;
- Existem 4 funcionalidades relevantes para as televisões: decodificação de sinal digital, desligamento automático, mute e tecla sap.

a) Modele as informações através de frames



b) Seria possível realizar a mesma modelagem usando somente redes semânticas? (justifique a sua resposta)

Não porque as funcionalidades decodificação de sinal digital, desligamento automático, mute e tecla sap são representadas como procedimentos, um conceito que não está presente nas redes semânticas.

## Lógica Nebulosa

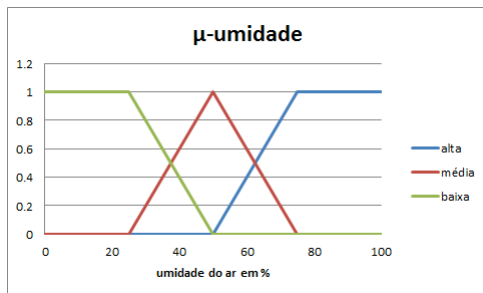
1. Dada as variáveis e as regras a seguir

Variáveis

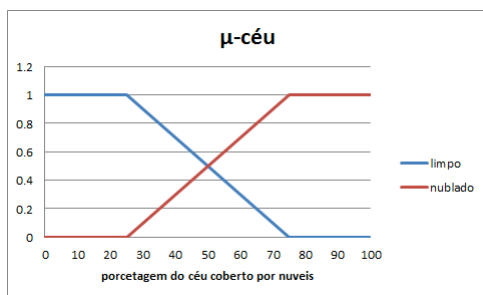
época do ano: [chuvosa, seca]



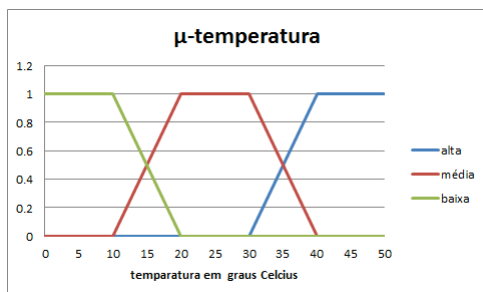
umidade: [baixa, alta, média]



céu: [limpo, nublado]



temperatura [baixa, média, alta]



*Regras:*

R1: Se época do ano= chuvosa e umidade = alta e céu = nublado então temperatura= baixa

R2: Se época do ano= chuvosa e umidade = baixa e céu = limpo então temperatura = média

R3: Se época do ano= seca e umidade = baixa ou céu = limpo então temperatura = alta

R4: Se época do ano = seca e umidade = alta e céu = nublado então temperatura = média

Para Época do ano = junho (mês 6), umidade = 50% e céu 35% coberto por nuvens, pede-se:

a) A fuzzyficação das variáveis

*Fuzzyficação*

$$\mu\text{-época do ano-seca}(6) = 0,5$$

$$\mu\text{-época do ano-chuvosa}(6) = 0,5$$

$$\mu\text{-umidade-baixa}(50\%) = 0$$

$$\mu\text{-umidade-média}(50\%) = 1$$

$$\mu\text{-umidade-alta}(50\%) = 0$$

$$\mu\text{-céu-limpo}(35\%) = 0,8$$

$$\mu\text{-céu-nublado}(35\%) = 0,2$$

b) Avaliação e das regras e a defuzzyficação do resultado para a descoberta da temperatura.

*Avaliação das regras – Inferência:*

$$R1: D^1 = \text{mim}[\mu\text{-época do ano-chuvosa}(6) = 0,5 ; \mu\text{-umidade-alta}(50\%) = 0; \mu\text{-céu -nublado}(35\%) = 0,2] = 0$$

(temperatura = baixa)

$$R2: D^2 = \text{mim}[\mu\text{-época do ano-chuvosa}(6) = 0,5 ; \mu\text{-umidade-baixa}(50\%) = 1; \mu\text{-céu -limpo}(35\%) = 0,8] = 0,5$$

(temperatura = média)

$$R3: D^3 = \text{mim}[\mu\text{-época do ano-seca}(6) = 0,5 ; \mu\text{-umidade-baixa}(50\%) = 1; \mu\text{-céu -limpo}(35\%) = 0,8] = 0,5$$

(temperatura = alta)

$$R4: D^4 = \text{mim}[\mu\text{-época do ano-seca}(6) = 0,5 ; \mu\text{-umidade-alta}(50\%) = 0; \mu\text{-céu -nublado}(35\%) = 0,2] = 0,0$$

(temperatura = média)

Defuzzyficação:



$$T = [(15 + 20 + 25 + 30 + 35 + 40) * 0,5 + (45 + 50) * 0,5] / (0,5 + 0,5 + 0,5 + 0,5 + 0,5 + 0,5 + 0,5 + 0,5) = [(165 * 0,5) + (95 * 0,5)] / 4 = (82,5 + 47,5) / 2,8 = 32,5$$

Para Época do ano = julho (mês 7), umidade = 60% e céu 50% coberto por nuvens, qual seria a temperatura?

a) A fuzzyficação das variáveis

### *Fuzzyficação*

$$\mu\text{-época do ano-seca}(7) = 0$$

$$\mu\text{-época do ano-chuvosa}(7) = 1$$

$$\mu\text{-umidade-baixa}(60\%) = 0$$

$$\mu\text{-umidade-média}(60\%) = 0,5$$

$$\mu\text{-umidade-alta}(60\%) = 0,5$$

$$\mu\text{-céu-limpo}(50\%) = 0,5$$

$$\mu\text{-céu-nublado}(50\%) = 0,5$$

b) Avaliação e das regras e a defuzzyficação do resultado para a descoberta da temperatura.

### *Avaliação das regras – Inferência:*

$$R1: D^1 = \min[\mu\text{-época do ano-chuvosa}(7) = 1 ; \mu\text{-umidade-alta}(60\%) = 0,5; \mu\text{-céu -nublado}(50\%) = 0,5] = 0,5 \text{ (temperatura = baixa)}$$

$$R2: D^2 = \min[\mu\text{-época do ano-chuvosa}(7) = 1 ; \mu\text{-umidade-baixa}(60\%) = 0,5; \mu\text{-céu -limpo}(50\%) = 0,5] = 0,5 \text{ (temperatura = média)}$$

$$R3: D^3 = \min[\mu\text{-época do ano-seca}(7) = 0 ; \mu\text{-umidade-baixa}(60\%) = 0,5; \mu\text{-céu -limpo}(50\%) = 0,5] = 0,5 \text{ (temperatura = alta)}$$

R4:  $D^4 = \min[\mu\text{-época do ano-seca}(7) = 0 ; \mu\text{-umidade-alta}(60\%) = 0,5; \mu\text{-céu -nublado}(50\%) = 0,5] = 0,0$   
 (temperatura = média)

Defuzzyficação:

$$T = [(5 + 10) * 0,5 + (15 + 20 + 25 + 30 + 35 + 40) * 0,5] / (0,5 + 0,5 + 0,5 + 0,5 + 0,5 + 0,5 + 0,5 + 0,5) = [(15 * 0,5) + (165 * 0,5)] / 4 = (90) / 4 = 22.5$$

## Redes Neurais

1. Dada a função  $y = f(a,b)$  representada pela tabela abaixo

Tabela

a	b	y
0	0	0
0	1	0
0	2	0
1	0	0
1	1	0
1	2	0
2	0	1
2	1	1
2	2	2

- a) Projete uma rede neural (composta por neurônios tipo McCulloch-Pitts com  $\alpha=1$ , treinada usando o aprendizado de Hebb e com a matriz de pesos inicialmente zerada) capaz de aprender como calcular a função representada pela tabela acima. Na resposta: (i) monte uma codificação para as entradas e saídas (ii) descreva a arquitetura (iii) mostre as matrizes geradas em cada passo do treinamento até se chegar a versão final da Matriz de sinapses;

Codificação

A	B	entrada	saída
0	0	[-1,-1,-1,-1]	[-1,-1]
0	1	[-1,-1,-1, 1]	[-1,-1]
0	2	[-1,-1, 1,-1]	[-1,-1]
1	0	[-1,-1, 1, 1]	[-1,-1]
1	1	[-1, 1,-1,-1]	[-1,-1]
1	2	[-1, 1,-1, 1]	[-1,-1]

2	0	[-1, 1, 1,-1]	[-1, 1]
2	1	[-1, 1, 1, 1]	[-1, 1]
2	2	[ 1,-1,-1,-1]	[1,- 1]

Arquitetura: quatro nós entrada e dois de saída

Matriz de pesos

Valores Iniciais

	F1	F2
A1	0	0
A2	0	0
A3	0	0
A4	0	0

Após o passo 1

	F1	F2
A1	1	1
A2	1	1
A3	1	1
A4	1	1

Após o Passo 2

	F1	F2
A1	2	2
A2	2	2
A3	2	2
A4	0	0

Após o Passo 3

	F1	F2
A1	3	3
A2	3	3
A3	1	1
A4	1	1

Após o Passo 4

	F1	F2
A1	4	4
A2	4	4
A3	0	0
A4	0	0

Após o Passo 5

	F1	F2
A1	5	5
A2	3	3
A3	1	1
A4	1	1

Após o Passo 6

	F1	F2
A1	6	6
A2	2	2
A3	2	2
A4	0	0

Após o Passo 7

	F1	F2
A1	7	5
A2	1	3
A3	1	3
A4	1	-1

Após o Passo 8

	F1	F2
A1	8	4
A2	0	4
A3	0	4
A4	0	0

### Após o Passo 9

	F1	F2
A1	9	3
A2	-1	5
A3	-1	5
A4	-1	1

- b) Usando a matriz de sinapses encontrada, calcule os valores gerados pela rede para cada uma das entradas;

Saída			Saída Convertida	
-6	-14		-1	-1
-8	-12		-1	-1
-8	-4		-1	-1
-10	-2		-1	-1
-8	-4		-1	-1
-10	-2		-1	-1
-10	6		-1	1
-12	8		-1	1
12	-8		1	-1

- c) O resultado obtido foi capaz de generalizar corretamente a função? (Justifique sua resposta)

Sim, isto ocorre porque o problema é linearmente separável (por duas retas) caso contrário, a rede provavelmente não seria capaz de generalizar o problema, podendo gerar resultados incorretos.

2. O que é o overfitting em redes neurais, como ele ocorre e como é possível evita-lo?

É um fenômeno que pode ocorrer durante o treinamento de uma rede neural caracterizado pelo alto desempenho da rede no conjunto de treinamento, aliada a falta de capacidade de atingir o mesmo desempenho em um conjunto de validação (composto exemplos diferentes dos processados durante o treinamento). Assim, a rede aprende apenas como classificar os exemplos que ela já conhece, mas não é capaz de generalizar tal conhecimento de forma que ele possa ser aplicado em exemplos desconhecidos. Uma das formas mais simples de se evitar o overfitting é através da validação cruzada, onde os exemplos são divididos em  $n$  subconjuntos (disjuntos). Desta forma, para se avaliar o desempenho da rede no  $i$ -ésimo subconjunto todos os demais subconjuntos são usados durante o treinamento da rede.

## Raciocínio Probabilístico

1. Para chegar ao seu trabalho Joana usa o metrô para ir do seu apartamento a estação que fica a uma quadra do escritório aonde trabalha. Há dias em que Joana pode ser atrasar no caminho para o seu escritório (se o metrô sofre algum tipo de problema). Problemas no metrô ocorrem devido a dois fatores: (1) número excessivo de passageiros; (2) falha na circulação de um dos trens da linha. Assim, dadas as variáveis e probabilidades condicionais seguir:

### *Variáveis*

Joana se atrasou:  $P(JA)$

Problemas de circulação no metrô:  $P(PM)$

Passageiros em excesso no trem:  $P(PE) = 0.7$

Falha na circulação de um dos trens:  $P(FT) = 0.1$

### *Probabilidades Condicionais*

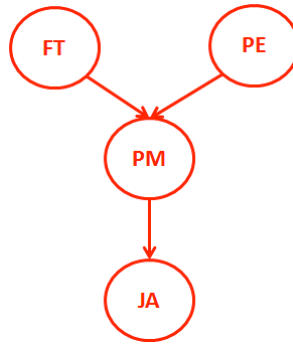
Joana se atrasou (JA)

PM	P(JA)
V	0,8
F	0,1

Problemas de circulação no metrô (PM)

PE	TF	P(PM)
V	V	0.8
V	V	0.3
F	V	0.2
F	F	0.1

- a) Monte uma rede de dependência das entre as variáveis acima descritas.



b) Qual a probabilidade da seguinte conjunção de eventos:

- Joana se atrasar
- Houve uma falha de circulação em um dos trens do metrô
- Não há passageiros em excesso no metrô
- O metrô não teve problemas

$$P(\text{JA}, \text{FT}, \neg\text{PE}, \neg\text{PM}) = P(\text{JA} | \neg\text{PM}) * P(\text{FT}) * P(\neg\text{PE}) * P(\neg\text{PM} | \text{FT}, \neg\text{PE}) = 0.1 * 0.1 * 0.3 * 0.8 = 0.0024$$

c) Dado que Joana se atrasou qual a probabilidade ter ocorrido um problema no metro?

$$P(\text{PM} | \text{JA}) = P(\text{PM}, \text{JA}) / P(\text{JA})$$

$$P(\text{JA}) = P(\text{JA}, \text{PM}, \text{PE}, \text{FL}) + P(\text{JA}, \text{PM}, \text{PE}, \neg\text{FL}) + P(\text{JA}, \text{PM}, \neg\text{PE}, \text{FL}) + P(\text{JA}, \text{PM}, \neg\text{PE}, \neg\text{FL}) + P(\text{JA}, \neg\text{PM}, \text{PE}, \text{FL}) + P(\text{JA}, \neg\text{PM}, \text{PE}, \neg\text{FL}) + P(\text{JA}, \neg\text{PM}, \neg\text{PE}, \text{FL}) + P(\text{JA}, \neg\text{PM}, \neg\text{PE}, \neg\text{FL}) = 0.1668 \text{ (ver como chegar nesta resposta nas próximas linhas)}$$

$$P(\text{JA}, \text{PM}, \text{PE}, \text{FL}) = P(\text{JA} | \text{PM}) * P(\text{PM} | \text{PE}, \text{FL}) * P(\text{PE}) * P(\text{FL}) = 0.6 * 0.8 * 0.7 * 0.1 = 0.0336$$

$$P(\text{JA}, \text{PM}, \text{PE}, \neg\text{FL}) = P(\text{JA} | \text{PM}) * P(\text{PM} | \text{PE}, \neg\text{FL}) * P(\text{PE}) * P(\neg\text{FL}) = 0.6 * 0.3 * 0.7 * 0.9 = 0.1134$$

$$P(\text{JA}, \text{PM}, \neg\text{PE}, \text{FL}) = P(\text{JA} | \text{PM}) * P(\text{PM} | \neg\text{PE}, \text{FL}) * P(\neg\text{PE}) * P(\text{FL}) = 0.6 * 0.2 * 0.3 * 0.1 = 0.0036$$

$$P(\text{JA}, \text{PM}, \neg\text{PE}, \neg\text{FL}) = P(\text{JA} | \text{PM}) * P(\text{PM} | \neg\text{PE}, \neg\text{FL}) * P(\neg\text{PE}) * P(\neg\text{FL}) = 0.6 * 0.1 * 0.3 * 0.9 = 0.0162$$

$$P(JA, \neg PM, EP, FL) = 0$$

$$P(JA, \neg PM, EP, \neg FL) = 0$$

$$P(JA, \neg PM, \neg EP, FL) = 0$$

$$P(JA, \neg PM, \neg EP, \neg FL) = 0$$

Todas dão 0 já que  $P(JA | PM) = 0$

$$P(PM, JA) = P(JA, PM) = P(JA, PM, PE, FL) + P(JA, PM, PE, \neg FL) + P(JA, PM, \neg PE, FL) + P(JA, PM, \neg PE, \neg FL) = 0.1668$$

Caímos na mesma equação da primeira parte do exercício, por isso o cálculo de  $P(PM, JA)$  torna-se trivial.

$$\text{Resposta: } P(PM | JA) = P(PM, JA) / P(JA) = 0.1668 / 0.1668 = 1$$