

2º Teste (Modelo A) – Com consulta limitada –

D) [6val] Uma companhia aérea pretende renovar o fardamento dos seus assistentes de bordo. Para o efeito, especificou uma série de condições para o “design” das fardas. Em primeiro lugar, só é considerado como calçado botas, sapatos ou ténis. Como partes de cima foram seleccionadas fato, camisa, blusa ou casaco, e como partes de baixo novamente o fato e ainda calça, saia ou calção. O objectivo é obter um fardamento para homens e mulheres obtido a partir de combinações das peças anteriores, juntamente com outras restrições adicionais a detalhar. Após análise do problema decidiu-se utilizar a programação por conjuntos de resposta, no dialecto do CLINGO, para sugerir os fardamentos possíveis, utilizando como manequins profissionais o António, o Bernardo, a Catarina e a Diana. Para facilitar a sua resolução, o seguinte conhecimento já se encontra devidamente modelado:

calçado(botas). calçado(sapatos). calçado(ténis).	cima(fato). cima(camisa). cima(blusa). cima(casaco).	baixo(fato). baixo(calça). baixo(saia). baixo(calção).	homem(antónio). homem(bernardo). mulher(catarina). mulher(diana).
pessoa(X) :- homem(X). pessoa(X) :- mulher(X).	1 {coloca(X,Y):calçado(Y)} 1:- pessoa(X). 1 {coloca(X,Y):cima(Y)} :- pessoa(X). 1 {coloca(X,Y):baixo(Y)} 1 :- pessoa(X).		

Para responder às seguintes questões, pode utilizar predicados auxiliares, caso entenda necessário.

- Explique qual é o efeito da seguinte restrição de integridade, quando adicionada ao programa anterior:

$$:- \text{pessoa}(X), \text{not coloca}(X, \text{calção}).$$
- Especifique as regras/restrições necessárias para garantir que os homens não usam blusas nem saias.
- A companhia aérea pretende impor que quando alguém coloca fato tem de usar blusa ou camisa. Especifique as regras/restrições necessárias para representar esta imposição.
- Especifique as regras/restrições necessárias para obrigar a que todas as pessoas usem o mesmo calçado.
- Explique sucintamente o efeito da restrição abaixo e apresente uma fórmula em lógica de primeira ordem equivalente a essa restrição imposta pela companhia aérea:

$$\begin{aligned}
 &:- \text{not } r. \\
 &r :- \text{baixo}(P), \text{aux1}(P). \\
 &\text{aux1}(P) :- \text{baixo}(P), \text{not aux2}(P). \\
 &\text{aux2}(P) :- \text{pessoa}(X), \text{baixo}(P), \text{not coloca}(X, P).
 \end{aligned}$$

XX

II) [4val] Considere os seguintes atributos e respectivos valores possíveis:

$$x_1 \in \{\text{old}, \text{mid}, \text{new}\} \quad x_2 \in \{Y, N\}$$

e o seguinte conjunto de 7 exemplos a ser usados na construção de uma árvore de decisão usando o algoritmo DTL.

	x_1	x_2	Classificação
D ₁	old	Y	+
D ₂	old	Y	+
D ₃	old	Y	-
D ₄	old	N	-

	x_1	x_2	Classificação
D ₅	mid	N	-
D ₆	mid	Y	+
D ₇	mid	N	-

- Qual o ganho de informação (IG) de cada um dos atributos? Apresente os cálculos. Os seguintes valores de entropia poderão ajudar:

x	y	$H\left(\frac{x}{y}, 1 - \frac{x}{y}\right)$
1	1	0,00
1	2	1,00
1	3	0,92

x	y	$H\left(\frac{x}{y}, 1 - \frac{x}{y}\right)$
1	4	0,81
1	5	0,72
1	6	0,65

x	y	$H\left(\frac{x}{y}, 1 - \frac{x}{y}\right)$
1	7	0,59
1	8	0,54
2	5	0,97

x	y	$H\left(\frac{x}{y}, 1 - \frac{x}{y}\right)$
2	7	0,86
3	7	0,99
3	8	0,95

- Qual o atributo a ser escolhido como raiz da árvore? Justifique. Se necessário, desempate a favor do atributo de menor índice (e.g. x_1 vence sobre x_2).
- Apresente a árvore de decisão induzida pelo algoritmo DTL. Justifique e apresente os cálculos efectuados.

III) [1val] Para cada uma das seguintes frases em Lógica de Primeira Ordem da coluna da esquerda, indique a letra que corresponde à frase em língua natural na coluna da direita da qual ela é a melhor representação. Considere as seguintes interpretações: (1) “Aluno(x)” significa que “x é um aluno”; (2) “Quiz(x)” significa que “x é um quiz”; (3) “Nota100(x,y)” significa que “x teve 100 em y”.

1.	$\forall s \exists q \text{ Aluno}(s) \Rightarrow [\text{Quiz}(q) \wedge \text{Nota100}(s, q)]$	A	Todo o quiz teve algum aluno com 100
2.	$\exists q \forall s \text{ Quiz}(q) \wedge [\text{Aluno}(s) \Rightarrow \text{Nota100}(s, q)]$	B	Todo o aluno teve 100 nalgum quiz
3.	$\forall q \exists s \text{ Quiz}(q) \Rightarrow [\text{Aluno}(s) \wedge \text{Nota100}(s, q)]$	C	Todos os alunos tiveram 100 em todos os quizzes
4.	$\exists s \forall q \text{ Aluno}(s) \wedge [\text{Quiz}(q) \Rightarrow \text{Nota100}(s, q)]$	D	Algum aluno teve 100 nalgum quiz
5.	$\forall s \forall q [\text{Aluno}(s) \wedge \text{Quiz}(q)] \Rightarrow \text{Nota100}(s, q)$	E	Há um quizz em que todos os alunos tiveram 100
6.	$\exists s \exists q \text{ Aluno}(s) \wedge \text{Quiz}(q) \wedge \text{Nota100}(s, q)$	F	Há um aluno que teve 100 em todos os quizzes

XX

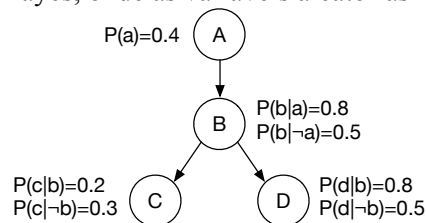
IV) [2val] Sejam **A**, **B**, **C** e **D** as seguintes acções na linguagem STRIPS:

Acção: A	Acção: B	Acção: C	Acção: D
Precondições: b	Precondições: c	Precondições: a	Precondições: b
Efeitos: a, ~b	Efeitos: a, b	Efeitos: b, c, ~a	Efeitos: d

- Apresente um plano de ordem parcial tal como gerado pelo algoritmo POP para atingir o estado em que **a** e **b** são verdadeiros, partindo do estado em que apenas **a** é verdadeiro. Indique uma linearização do plano.
- Suponha que o algoritmo GraphPlan consegue extrair uma solução no nível n de um grafo de planeamento. Esse plano é necessariamente ótimo? O seu tamanho (número de ações) é igual a n? Justifique.

XX

V) [5val] Considere a seguinte Rede de Bayes, onde as variáveis aleatórias A, B, C e D são todas booleanas.



Em cada uma das seguintes alíneas, calcule a probabilidade indicada, usando o método indicado.

- $P(\neg c)$ usando o método de inferência por enumeração.
- $P(b|\neg c)$ usando o método de inferência por enumeração.
- $P(b|\neg c)$ usando o método de inferência por eliminação de variáveis.

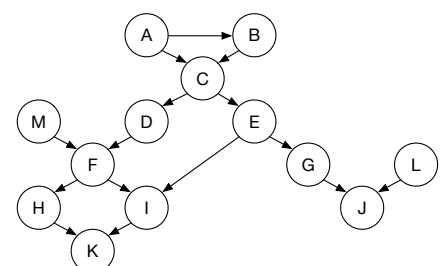
XX

VI) [2val] Para cada alínea, indique se ela é verdadeira ou falsa. Cada resposta correta vale 0.4val, cada resposta errada vale -0.4val. A pergunta tem uma cotação mínima de 0 valores.

- Numa Rede de Bayes, se A é independente de B, então A é independente de B dado C, onde C é uma variável diferente de A e B.
- Um perceptrão pode aprender e representar qualquer função Booleana.
- O ganho de informação no nó raiz de uma qualquer árvore de decisão é sempre maior ou igual do que o ganho de informação de qualquer outro nó dessa árvore.
- Duas árvores de decisão diferentes aprendidas com o mesmo conjunto de treino (e.g. aprendidas com algoritmos diferentes), irão atribuir necessariamente a mesma classificação a novos exemplos.
- A ordem pela qual os exemplos do conjunto de treino são processados pelo algoritmo de eliminação de candidatos não afecta o resultado final.

XX

VII) [Bónus: até 2val] Considere a Rede de Bayes à direita, onde as variáveis aleatórias são todas booleanas e as tabelas de probabilidade condicionada foram omitidas. Apresente uma expressão, o mais simples possível, que permita calcular a distribuição de probabilidade condicional $P(D|\neg b)$ em função dos valores das tabelas de probabilidade condicionada da Rede de Bayes. A resposta será avaliada em função da sua correção e simplicidade. Uma expressão desnecessariamente complexa, ainda que correta, terá uma classificação de 0 valores.



Nome:

Número:

I.a)

Impede a existência de modelos estáveis em que alguém não use calções, ou seja, só permite modelos estáveis em que todas as pessoas usem calções.

I.b)

```
:- homem(X), coloca(X,saia).
:- homem(X), coloca(X,blusa).
```

I.c)

```
ok(X) :- pessoa(X), coloca(X,camisa).
ok(X) :- pessoa(X), coloca(X,blusa).
:- pessoa(X), coloca(X,fato), not ok(X).
```

I.d)

```
:- pessoa(X1), pessoa(X2), calçado(Y1), calçado(Y2),
   coloca(X1,Y1), coloca(X2,Y2), X1 != X2, Y1 != Y2.
```

I.e)

Fórmula: $\exists x [baixo(x) \wedge \forall y [pessoa(y) \Rightarrow coloca(y, x)]]$

Explicação: **aux2** é verdadeiro para uma dada parte de baixo P quando essa parte de baixo não foi usada por alguma pessoa. **aux1** é verdadeiro quando **aux2** é falso, ou seja, quando uma dada parte de baixo é usada por todas as pessoas. O átomo **r** é verdadeiro quando existe alguma parte de baixo para o qual **aux1** seja verdadeiro i.e. **r** é verdadeiro quando alguma parte de baixo foi usada por todas as pessoas. A primeira regra (restrição de integridade) impede que **r** seja falso i.e. obriga a que **r** seja verdadeiro i.e. obriga a que exista uma parte de baixo (*x* na fórmula) que seja usada por todas as pessoas (*y* na fórmula).

II.a) $IG(x_1) = 0,024$ $IG(x_2) = 0,527$

Cálculos:

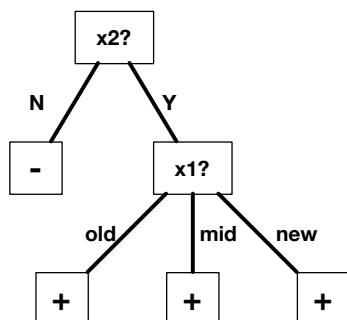
$$IG(x_1) = H\left(\frac{3}{7}, \frac{4}{7}\right) - \left[\frac{4}{7} \cdot H\left(\frac{2}{4}, \frac{2}{4}\right) + \frac{3}{7} \cdot H\left(\frac{1}{3}, \frac{2}{3}\right)\right] = 0,99 - \left(\frac{4}{7} \cdot 1 + \frac{3}{7} \cdot 0,92\right) = 0,024$$

$$IG(x_2) = H\left(\frac{3}{7}, \frac{4}{7}\right) - \left[\frac{4}{7} \cdot H\left(\frac{3}{4}, \frac{1}{4}\right) + \frac{3}{7} \cdot H\left(\frac{3}{3}, \frac{0}{3}\right)\right] = 0,99 - \left(\frac{4}{7} \cdot 0,81 + \frac{3}{7} \cdot 0\right) = 0,527$$

II.b) Atributo: x_2

Justificação: Porque tem o maior ganho de informação

II.c) Árvore:



Justificação/cálculos:

Com $x_2 = N$, todos os exemplos são classificados com “-” pelo que será um nó terminal com valor “-”. No caso de $x_2 = Y$ é necessário testar o atributo x_1 . Nesse teste, no caso de $x_1 = Old$, nem todos os exemplos concordam, mas como não há mais atributos, o nó é terminal com valor igual à moda i.e. “+”. Com $x_1 = Mid$ todos os exemplos (no caso apenas um) são classificados com “+” pelo que será um nó terminal com esse valor. Com $x_1 = New$ não há exemplos pelo que o nó é terminal com valor igual à moda do conjunto de exemplos do nó pai (x_1) i.e. “+”.

III	1. B	2. E	3. A	4. F	5. C	6. D
IV.a) Plano: (Nota: existem outras alternativas.)	<div>Start</div> <div>a</div> <div>↓</div> <div>a</div> <div>C</div> <div>~a b c</div> <div>↙</div> <div>c</div> <div>B</div> <div>a b</div> <div>↓ ↓</div> <div>a b</div> <div>Finish</div>	<div>Restrições:</div> <div>Start < Finish</div> <div>Start < B</div> <div>B < Finish</div> <div>Start < C</div> <div>C < Finish</div> <div>C < B</div> <div>Linearização: C→B</div>				
IV.b) Óptimo [SIM/NÃO]: NÃO Tamanho igual a n [SIM/NÃO]: NÃO Justificação: O resultado do GraphPlan é um plano de ordem parcial, podendo conter mais do que uma ação por nível – logo, o número de acções pode ser superior a n. Mais, como podem existir vários planos parciais no mesmo grafo de planeamento, potencialmente com um número de acções diferente, e o algoritmo de extração da solução tem não-determinismo na escolha de acções, não dando portanto garantias de qual dos planos existentes vai escolher, não temos a garantia que o plano escolhido seja óptimo no número de ações, entre os planos de nível em causa. Acresce que o plano óptimo pode nem sequer existir no nível n, podendo haver um plano disponível num nível superior a n no grafo de planeamento com menos acções do que todos os planos disponíveis no nível n.						
V.a) Nota: o método de inferência por enumeração não tira partido das simplificações que se podem fazer com base na topologia da Rede de Bayes. Nomeadamente, no caso desta alínea, o método não ignora a variável D, apesar dela ser irrelevante (por não ser parte do conjunto composto pela variável de consulta, variável de evidência, e seus ascendentes). Juntamente com o facto de muitas vezes obrigar a repetições de cálculos, é uma das razões pelas quais o método de inferência por enumeração é ineficiente.						
$\begin{aligned} P(\neg c) &= \sum_A \sum_B \sum_D P(A, B, \neg c, D) = \\ &= \sum_A \sum_B \sum_D P(A) P(B A) P(\neg c B) P(D B) = \\ &= \sum_A P(A) \sum_B P(B A) P(\neg c B) \sum_D P(D B) = \\ &= 0,4 \cdot [0,8 \cdot 0,8 \cdot (0,8 + 0,2) + 0,2 \cdot 0,7 \cdot (0,5 + 0,5)] + 0,6 \cdot [0,5 \cdot 0,8 \cdot (0,8 + 0,2) + 0,5 \cdot 0,7 \cdot (0,5 + 0,5)] = \\ &= 0,762 = 76,2\% \end{aligned}$						
V.b) Nota: a nota da alínea anterior também se aplica nesta alínea.						
$P(b \neg c) = \alpha P(b, \neg c) = \alpha \sum_A \sum_D P(A, b, \neg c, D)$						
$\begin{aligned} \sum_A \sum_D P(A, B, \neg c, D) &= \sum_A \sum_D P(A) P(B A) P(\neg c B) P(D B) = \\ &= P(\neg c B) \sum_A P(A) P(B A) \sum_D P(D B) = \end{aligned}$						
$\begin{aligned} \mathbf{B} = true: \quad &0,8 \cdot [0,4 \cdot 0,8 \cdot (0,8 + 0,2) + 0,6 \cdot 0,5 \cdot (0,8 + 0,2)] = 0,496 \\ \mathbf{B} = false: \quad &0,7 \cdot [0,4 \cdot 0,2 \cdot (0,5 + 0,5) + 0,6 \cdot 0,5 \cdot (0,5 + 0,5)] = 0,266 \\ P(b \neg c) &= \frac{0,496}{0,496 + 0,266} \approx 0,6509 = 65,09\% \end{aligned}$						

Nome:

Número:

V.c)

$$\begin{aligned}
 P(b|\neg c) &= \\
 &= \alpha P(\neg c|B) \sum_A P(A) P(B|A) \sum_D P(D|B) \\
 &= \alpha P(\neg c|B) \sum_A P(A) P(B|A) \sum_D f_D(B,D) \\
 &= \alpha P(\neg c|B) \sum_A P(A) P(B|A) f_{\bar{D}}(B) \\
 &= \alpha P(\neg c|B) \sum_A P(A) f_B(A,B) f_{\bar{D}}(B) \\
 &= \alpha P(\neg c|B) \sum_A f_A(A) f_B(A,B) f_{\bar{D}}(B) \\
 &= \alpha P(\neg c|B) \sum_A f_{AB\bar{D}}(A,B) \\
 &= \alpha P(\neg c|B) f_{\bar{AB}\bar{D}}(B) \\
 &= \alpha f_C(B) f_{\bar{AB}\bar{D}}(B) \\
 &= \alpha f_{\bar{ABC}\bar{D}}(B) \\
 &= \frac{0,496}{0,496 + 266} \approx 0,6509 = 65,09\%
 \end{aligned}$$

$$f_D$$

B	D	$f_D(B,D)$
t	t	0,8
t	f	0,2
f	t	0,5
f	f	0,5

$$f_{\bar{D}}$$

B	$f_{\bar{D}}(B)$
t	1
f	1

$$f_B$$

A	B	$f_B(A,B)$
t	t	0,8
t	f	0,2
f	t	0,5
f	f	0,5

$$f_A$$

A	$f_A(A)$
t	0,4
f	0,6

$$f_{AB\bar{D}}$$

A	B	$f_{AB\bar{D}}(A,B)$
t	t	$0,8 \times 0,4 = 0,32$
t	f	$0,2 \times 0,4 = 0,08$
f	t	$0,5 \times 0,6 = 0,3$
f	f	$0,5 \times 0,6 = 0,3$

$$f_{\bar{AB}\bar{D}}$$

B	$f_{\bar{AB}\bar{D}}(B)$
t	$0,32 + 0,3 = 0,62$
f	$0,08 + 0,3 = 0,38$

$$f_C$$

B	$f_C(B)$
t	0,8
f	0,7

$$f_{\bar{ABC}\bar{D}}$$

B	$f_{\bar{ABC}\bar{D}}(B)$
t	$0,62 \times 0,8 = 0,496$
f	$0,38 \times 0,7 = 0,266$

VI

i. FALSA

ii. FALSA

iii. FALSA

iv. FALSA

v. VERDADEIRA

VII

$$P(D|\neg b) = \frac{\sum_A P(A) P(\neg b|A) \sum_C P(C|A, \neg b) P(D|C)}{\sum_A P(A) P(\neg b|A)}$$