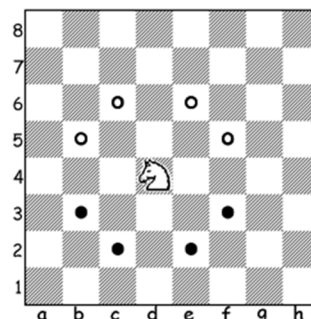


**Nota: Responder a cada questão (1, 2, 3 e 4) em folhas de exame separadas.**

1. [4 valores] Num tabuleiro de xadrez, um cavalo movimenta-se dando “saltos em L”, de acordo com o exemplo da figura: o cavalo na posição d4 pode deslocar-se, com um só salto, para as posições b3, b5, c2, c6, e2, e6, f3 ou f5. Para simplificar, ao cavalo considerado serão permitidos apenas saltos “para cima”: por exemplo, da posição d4 só será possível ir para b5, c6, e6 ou f5.



- a) Usando a estratégia de pesquisa **primeiro em profundidade**, apresente uma árvore de pesquisa e identifique a solução encontrada para que o cavalo se desloque da posição a1 para a posição f8.
- b) Neste cenário, sendo o objectivo do cavalo ir de um ponto de partida até um ponto destino indicado, quais das estratégias de pesquisa sistemática são completas? Justifique.

Todas.

...

- c) Considere as seguintes funções heurísticas possíveis que procuram estimar o número de saltos em falta para, a partir da posição atual  $i = x_i y_i$ , alcançar a posição final  $f = x_f y_f$ :

$$h_1 = |x_f - x_i| + |y_f - y_i| \quad (\text{distância Manhattan})$$

$$h_2 = \max(|x_f - x_i|, |y_f - y_i|) \quad (\text{distância Chebyshev})$$

$$h_3 = \max(|x_f - x_i|, |y_f - y_i|)/2$$

$$h_4 = \min(|x_f - x_i|, |y_f - y_i|)$$

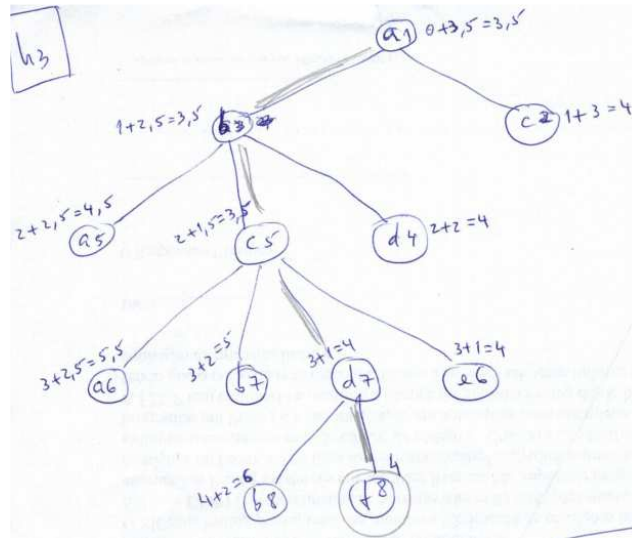
$$h_5 = \min(|x_f - x_i|, |y_f - y_i|)/2$$

Quais destas heurísticas são admissíveis? Justifique. Das admissíveis, qual é a melhor? Porquê?

h3 e h5

...

- d) Usando a função heurística identificada na resposta à alínea anterior, apresente uma árvore de pesquisa obtida pela aplicação da estratégia de pesquisa **A\***, de modo a deslocar o cavalo da posição a1 para a posição f8. Junto a cada nó da árvore, indique o valor dos componentes da função de custo ( $f=g+h$ ). Identifique a solução encontrada.



2. [4 valores] O problema de colorir um mapa consiste em atribuir cores diferentes a países que possuam fronteira entre si. A figura seguinte representa o mapa a colorir, constituído por 5 países (P1..P5).

Existem 3 cores disponíveis: azul (A), verde (V) e branco (B).

Pretende-se aplicar **Algoritmos Genéticos** na resolução deste problema. Suponha a existência de uma população inicial de 5 indivíduos com a seguinte informação: *corP1*, *corP2*, *corP3*, *corP4*, *corP5*

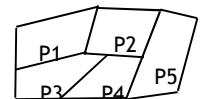
I1: B,V,B,A,V

I2: A,V,V,A,B

I3: B,V,A,V,V

I4: B,B,B,B,V

I5: B,B,V,B,V



- a) Proponha uma estrutura para a representação do indivíduo, explicando. Exemplifique com a representação do indivíduo I1 da população inicial.

Indivíduo: 10 bits, representam as cores dos 5 países. Cada conjunto de 2 bits representa a cor de um país. 00-azul, 01-verde, 10-branco

Indivíduo I1 = 1001100001

- b) Proponha uma função de adaptação (descrição textual). Calcule os valores de adaptação dos indivíduos da população inicial.

numErradas: nº de fronteiras erradas (uma fronteira está errada quando os dois países a ela adjacente possuem a mesma cor).

Para converter em problema de maximização, considera-se

função de adaptação = 7-numErradas (7 é o nº de fronteiras existentes)

Indivíduo I1: 7-2 = 5

Indivíduo I2: 7-1 = 6

Indivíduo I3: 7-3 = 4

Indivíduo I4: 7-5 = 2

Indivíduo I5: 7-2 = 5

- c) No processo de seleção dos indivíduos a utilizar na formação da geração seguinte, é usada uma política elitista (só para o melhor). Considere que foram gerados os seguintes números aleatórios (entre 0 e 1): 0.7 / 0.35 / 0.15 / 0.81. Apresente o resultado deste processo de seleção. Explique.

soma dos valores de adaptação = 22

probabilidade(I2) =  $6/22 = 0.272 \rightarrow$  intervalo ]0 ; 0.272]

probabilidade(I1) =  $5/22 = 0.227 \rightarrow$  intervalo ]0.272 ; 0.5]

probabilidade(I5) =  $5/22 = 0.227 \rightarrow$  intervalo ]0.5 ; 0.727]

probabilidade(I3) =  $4/22 = 0.182 \rightarrow$  intervalo ]0.727 ; 0.9]

probabilidade(I4) =  $2/22 = 0.091 \rightarrow$  intervalo ]0.91 ; 1]

Resultado da seleção:

I2 (elitismo, é o mais apto) 0 0 0 1 0 1 0 0 1 0

I5 (valor aleatório=0.7) 1 0 1 0 0 1 1 0 0 1

I1 (valor aleatório=0.35) 1 0 0 1 1 0 0 0 0 1

I2 (valor aleatório=0.15) 0 0 0 1 0 1 0 0 1 0

I3 (valor aleatório=0.81) 1 0 0 1 0 0 0 1 0 1

- d) Calcule a 2ª geração da população, explicando todas as suas opções. Sugira uma estratégia de cruzamento. A probabilidade de cruzamento é 75% e foram gerados os números aleatórios: 0.81 / 0.41 / 0.24 / 0.88. A probabilidade de mutação é 2% e só no 15º número aleatório surgiu um inferior a 0.02.

Estratégia de cruzamento: (por exemplo) cruzar os indivíduos num ponto de corte múltiplo de 2 (para não provocar alocações inválidas). Consideremos o ponto de corte entre os bits 4 e 5

1º filho:

- 4 primeiros bits do 1ºpai, 6 últimos bits do 2ºpai

2º filho:

- 4 primeiros bits do 2ºpai, 6 últimos bits do 1ºpai

I2: mantém (elitismo)

I5: mantém (valor aleatório=0.81 > 0.75)

I1: cruza (valor aleatório=0.41 < 0.75)

I2: cruza (valor aleatório=0.24 < 0.75)

I3: mantém (valor aleatório=0.88 > 0.75)

Após cruzamento:

I2 : 0 0 0 1 0 1 0 0 1 0

I5 : 1 0 1 0 0 1 1 0 0 1 mutação: 5ºbit mudaria, mas fica inválido, então não muda

I12: 1 0 0 1 0 1 0 0 1 0

I21: 0 0 0 1 1 0 0 0 0 1

I3: 1 0 0 1 0 0 0 1 0 1

Barbatana	Cauda	Corpo	Peixe
larga	grande	gordo	vermelho
larga	pequena	esguio	vermelho

Prova com consulta. Duração: 2h30m.

Exame da Época Normal

3. [4 valores] Numa loja de animais, um robô está encarregado da alimentação dos peixes, que é diferente para os peixes vermelhos e os peixes azuis. O robô irá aprender a diferenciar estas duas espécies de peixes através das partes do corpo, usando o conjunto de exemplos da tabela. (Nota: nas alíneas seguintes, atenda à ocorrência de um valor desconhecido.)

larga	grande	gordo	vermelho
finá	pequena	esguio	vermelho
finá	pequena	esguio	azul
finá	--	gordo	azul
larga	pequena	gordo	azul
finá	pequena	gordo	azul

- a) Calcule os valores da informação média para os exemplos considerados.

$\text{info\_media} = -4/8 * \log_2(4/8) - 4/8 * \log_2(4/8) = 0.5 + 0.5 = 1$  (quando considerar os atributos Barbatana e Corpo)

$\text{info\_media} = -4/7 * \log_2(4/7) - 3/7 * \log_2(3/7) = 0.461 + 0.524 = 0.985$  (quando considerar o atributo Cauda)

- a) Que atributo escolheria para raiz da árvore de decisão a construir, usando o critério da razão do ganho? Sabe-se que  $\text{entropia}(\text{Barbatana})=0.811$ ,  $\text{entropia}(\text{Cauda})=0.6$  e  $\text{entropia}(\text{Corpo})=0.95$ . Apresente todos os cálculos.

Barbatana

$$\text{GanhoInfo} = 1 - 0.811 = 0.188$$

$$\text{InfoSeparacao} = -4/8 * \log_2(4/8) - 4/8 * \log_2(4/8) = 0.5 + 0.5 = 1$$

$$\text{RazaoGanho} = 0.188 / 1 = 0.188$$

Cauda

$$\text{GanhoInfo} = 7/8 * (0.985 - 0.6) = 0.33$$

$$\text{InfoSeparacao} = -2/8 * \log_2(2/8) - 5/8 * \log_2(5/8) - 1/8 * \log_2(1/8) = 1.298$$

$$\text{RazaoGanho} = 0.33 / 0.984 = 0.26$$

Corpo

$$\text{GanhoInfo} = 1 - 0.95 = 0.05$$

$$\text{InfoSeparacao} = -3/8 * \log_2(3/8) - 5/8 * \log_2(5/8) = 0.954$$

$$\text{RazaoGanho} = 0.05 / 0.954 = 0.05$$

O atributo que maximiza a razão do ganho é "Cauda".

- b) Construa a árvore de decisão, considerando que a razão do erro em qualquer folha deve ser menor ou igual a 0.4. Apresente todos os cálculos e/ou explique convenientemente a sua resposta.

Por análise da amostra, o atributo Corpo é mais discriminativo que Barbatana (pois Corpo=gordo, tem entropia 0 Barbatana=larga tem entropia 1).

Por cálculos (não necessário se explicação prévia):

$$\text{InfoMedia} = -2/5 * \log_2(2/5) - 3/5 * \log_2(3/5) = 0.97$$

Barbatana

$$\text{Entropia} = -3/5 * [1/3 * \log_2(1/3) - 2/3 * \log_2(2/3)] + 2/5 * [1/2 * \log_2(1/2) - 1/2 * \log_2(1/2)] = 0.6$$

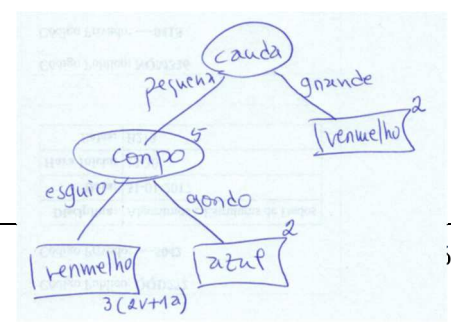
$$\text{GanhoInfo} = 0.97 - 0.6 = 0.37$$

$$\text{InfoSeparacao} = -3/5 * \log_2(3/5) - 2/5 * \log_2(2/5) = 0.97$$

$$\text{RazaoGanho} = 0.37 / 0.97 = 0.387$$

Corpo

$$\text{Entropia} = -3/5 * [2/3 * \log_2(2/3) - 1/3 * \log_2(1/3)] + 2/5 * [0 - 1 * \log_2(1)] = 0.344$$



$$\text{GanhoInfo} = 0.97 - 0.344 = 0.626$$

$$\text{InfoSeparacao} = -3/5 \cdot \log_2(3/5) - 2/5 \cdot \log_2(2/5) = 0.97$$

$$\text{RazaoGanho} = 0.626 / 0.97 = 0.65$$

$$\text{Razao Erro} = (e+1)/(n+2) = (1+1)/(3+2) = 0.4$$

4. [8 valores] Responda a seis (6) das seguintes sete (7) questões (cada uma em 5-10 linhas).

- a) O algoritmo de pesquisa heurística IDA\* usa uma pesquisa por aprofundamento iterativo na qual o limite de expansão é dado pelo custo e não pela profundidade. Como deve ser incrementado esse custo limite, de modo a que a solução ótima seja encontrada?
- b) Um estudante preparou-se para um exame, o que é indicativo de que passará (Crença=0.6); a mesma conclusão é suportada (Crença=0.7) pelo facto de que gosta da unidade curricular respetiva. Por outro lado, passou mal a noite, o que por si só pode indicar que vai chumbar no exame (Crença=0.3). Segundo o modelo de Dempster-Shafer, qual é o intervalo de confiança da aprovação no exame?

Evidências: preparou-se e gosta

		gosta	
		{passa}=0.7	{Θ}=0.3
preparou-se	{passa}=0.6	{passa}=0.42	{passa}=0.18
	{Θ}=0.4	{passa}=0.28	{Θ}=0.12

$$\text{Crença em } \{\text{passa}\} = 0.42 + 0.28 + 0.18 = 0.88$$

Evidência: dormiu mal

		dormiu mal	
		{~passa}=0.3	{Θ}=0.7
preparou-se e gosta	{passa}=0.88	{ }=0.264	{passa}=0.616
	{Θ}=0.12	{~passa}=0.036	{Θ}=0.084

$$\text{Crença em } \{\text{passa}\} = 0.616 / (1 - 0.264) = 0.616 / 0.736 = 0.837$$

$$\text{Crença em } \{\sim\text{passa}\} = 0.036 / (1 - 0.264) = 0.036 / 0.736 = 0.049$$

$$\text{Intervalo confiança em } \{\text{passa}\} = [0.837; 0.951]$$

- c) No algoritmo de otimização por “arrefecimento simulado”, relacione a probabilidade de escolha de um estado sucessor com a temperatura e com a diferença de valores entre o estado atual e esse estado.
- d) Os cortes alfa-beta, quando utilizados no algoritmo de pesquisa adversarial minimax, podem permitir reduzir o número de nós examinados de  $O(b^m)$  para  $O(b^{m/2})$ , sendo  $b$  o fator de ramificação e  $m$  a profundidade máxima da árvore de pesquisa. Explique em que circunstâncias é que isto acontece.
- e) Explique qual é o compromisso subjacente à utilização da fórmula “Naïve Bayes”.

- f) “Duas laranjas mais quatro peras são seis frutos.” Esboce uma DCG para interpretar frases deste tipo, em que se pretende validar operações sobre elementos cujo resultado é uma superclasse das parcelas. Indique que tipo de informação será necessário ter na base de conhecimento.
- g) Explique o conceito de *função linearmente separável* e relacione-o com a fórmula de cálculo da saída de um perceptrão.