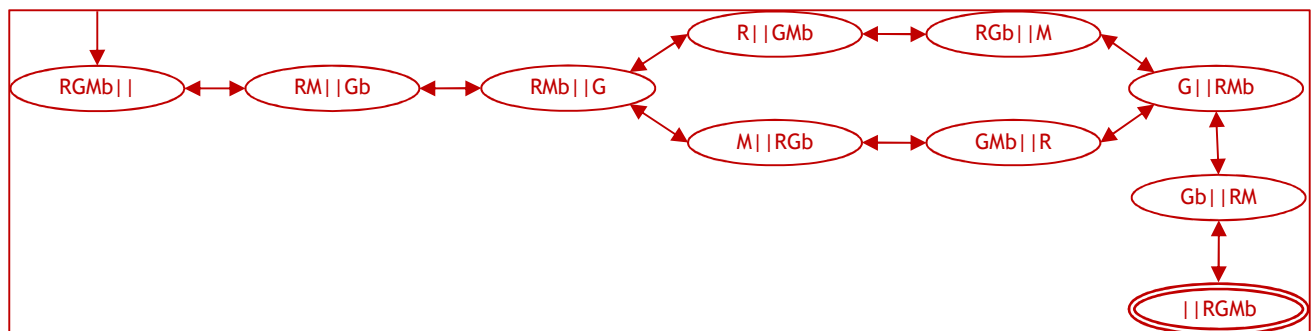


Nota: Responder a cada questão (1, 2, 3 e 4) em folhas de exame separadas.

1. [4 valores] Um camponês tem de atravessar um rio com uma raposa, uma galinha e uma cesta de milho. Tem um pequeno barco no qual pode transportar consigo, em qualquer direção, no máximo uma coisa de cada vez, mas não pode abandonar a raposa com a galinha, nem esta com a cesta de milho.

- a) Desenhe o espaço de estados com os estados legais do problema e transições entre eles. Represente o estado inicial da forma **RGMb||**, indicando que a raposa (R), a galinha (G), o milho (M) e o barco (b) estão na margem inicial do rio. Outros exemplos: o camponês transportou no barco a galinha para a outra margem (**RM|Gb**); o estado final: **||RGMb**.



- b) Sugira uma função heurística admissível (não trivial, i.e., $h=0$ não conta) que possa ser aplicada a este problema para avaliar cada estado, pretendendo-se minimizar o número de travessias. Mostre o valor da sua função para os estados identificados no enunciado da alínea anterior.

$h = 2n$ se b na margem esquerda

$h = 2n-1$ se b na margem direita

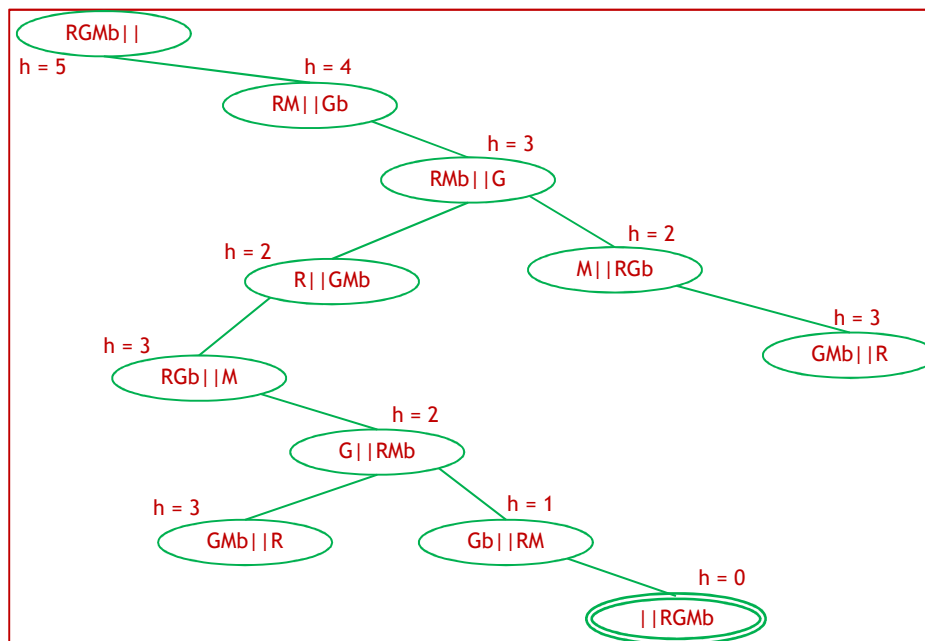
onde n é o número de “coisas” que ainda estão na margem inicial. A ideia é que para transferir cada coisa para a outra margem são necessários 2 passos: trazer o barco e ir com a coisa. O custo poderá ser maior caso haja necessidade de trazer de volta um dos objectos, “desfazendo” parte da tarefa.

$h(\text{RGMb}||) = 5$

$h(\text{RM}|\text{Gb}) = 4$

$h(||\text{RGMb}) = 0$

- c) Desenhe a árvore de pesquisa obtida pela aplicação da estratégia de **pesquisa gulosa (greedy)**, usando a função heurística definida na alínea anterior. Assinale, junto de cada nó da árvore, o valor da função heurística, e apresente a solução encontrada. Pode evitar representar transições que consubstanciem ciclos.



2. [4 valores] Pretende-se otimizar o plano de produção numa fábrica onde existem 3 máquinas: M1, M2, M3. A produção inclui o fabrico das peças A, B, C, D e E, que requerem o uso de máquinas específicas e têm uma duração (indicadas na tabela). Uma máquina só pode produzir uma peça de cada vez. Pretende-se determinar a alocação das diferentes peças às máquinas, de forma a minimizar o tempo total de produção.

Prod	Máquina	Duração
A	M1 ou M2	10
B	M1 ou M3	7
C	M2 ou M3	11
D	M2 ou M3	12
E	M1 ou M2 ou M3	8

- a) Suponha que partimos da alocação $S_1 = \{A-M1, B-M1, C-M2, D-M2, E-M1\}$. Calcule o custo desta alocação, $C(S_1)$.

$$C(S_1) = \max(10+7+8; 11+12) = \max(25; 23) = 25$$

- b) Considere uma função de vizinhança (para geração dos sucessores) que apenas altera a máquina de um dos produtos. Identifique todos os sucessores de S_1 . Segundo a variante “steepest ascent” do algoritmo “hill-climbing”, qual sucessor de S_1 será escolhido? Justifique com cálculos.

$$S_2 = \{A-M2, B-M1, C-M2, D-M2, E-M1\}; C(S_2) = 33$$

$$S_3 = \{A-M1, B-M3, C-M2, D-M2, E-M1\}; C(S_3) = 23$$

$$S_4 = \{A-M1, B-M1, C-M3, D-M2, E-M1\}; C(S_4) = 25$$

$$S_5 = \{A-M1, B-M1, C-M2, D-M3, E-M1\}; C(S_5) = 25$$

$$S_6 = \{A-M1, B-M1, C-M2, D-M2, E-M2\}; C(S_6) = 31$$

$$S_7 = \{A-M1, B-M1, C-M2, D-M2, E-M3\}; C(S_6) = 23$$

Será escolhido $S_3 = \{A-M1, B-M3, C-M2, D-M2, E-M1\}$ ou $S_7 = \{A-M1, B-M1, C-M2, D-M2, E-M3\}$.

- c) Considere que a função de vizinhança gera sucessores de forma determinística, alterando primeiro a máquina de A, depois de B, depois de C, depois de D e finalmente de E. Usando um valor de temperatura $t=10$, determine qual o sucessor de S_1 adotado pelo algoritmo de **arrefecimento simulado**, justificando com cálculos. Utilize, se necessário, os seguintes valores aleatórios gerados: 0.55, 0.89, 0.12, 0.45, 0.11, 0.66, 0.75.

$$p(S_2) = e^{-(\Delta E/t)} = e^{-(25-33)/10} = e^{(-0.8)} = 0.449$$

$S_2 \rightarrow$ rejeitado (valor aleatório = $0.55 > 0.449$)

$S_3 \rightarrow$ aceite, pois é um estado melhor do que S_1 .

3. [4 valores] Chegado o Verão, há que ter cuidado com o sol. A tabela ao lado mostra o que aconteceu a cada uma de 8 pessoas, com base nos seus dados fisionómicos, e sabendo se usou ou não protetor solar.

Name	Hair	Height	Weight	Lotion	Result
Sarah	blonde	average	light	no	sunburned
Dana	blonde	tall	average	yes	none
Alex	brown	short	average	yes	none
Annie	blonde	short	average	no	sunburned
Emily	red	average	heavy	no	sunburned
Pete	brown	tall	heavy	no	none
John	brown	average	heavy	no	none
Katie	blonde	short	light	yes	none

- a) Calcule a informação média relativamente à classificação desta coleção.

$$E(S) = -3/8 \log_2(3/8) - 5/8 \log_2(5/8) = 0.954$$

- b) Usando o algoritmo ID3, determine que atributo deve ser colocado na raiz da árvore de decisão, justificando com cálculos.

$$E(\text{res}/\text{hair}) = 4/8 [-2/4 \log_2(2/4) - 2/4 \log_2(2/4)] + 3/8 [-3/3 \log_2(3/3)] + 1/8 [-1/1 \log_2(1/1)] = 0.5$$

$$E(\text{res}/\text{height}) = 3/8 [-2/3 \log_2(2/3) - 1/3 \log_2(1/3)] + 2/8 [-2/2 \log_2(2/2)] + 3/8 [-2/3 \log_2(2/3) - 1/3 \log_2(1/3)] = 0.689$$

$$E(\text{res}/\text{weight}) = 2/8 [-1/2 \log_2(1/2) - 1/2 \log_2(1/2)] + 3/8 [-2/3 \log_2(2/3) - 1/3 \log_2(1/3)] + 3/8 [-1/3 \log_2(1/3) - 2/3 \log_2(2/3)] = 0.939$$

$$E(\text{res}/\text{lotion}) = 5/8 [-3/5 \log_2(3/5) - 2/5 \log_2(2/5)] + 3/8 [-3/3 \log_2(3/3)] = 0.607$$

O atributo escolhido é "hair".

- c) O algoritmo C4.5 usa o critério da razão do ganho para construir a árvore de decisão. Neste caso, determine qual o atributo que deve ser colocado na raiz da árvore de decisão, justificando com cálculos.

$$\text{infosep}(\text{hair}) = -4/8 \log_2(4/8) - 3/8 \log_2(3/8) - 1/8 \log_2(1/8) = 1.406$$

$$\text{infosep}(\text{height}) = -3/8 \log_2(3/8) - 2/8 \log_2(2/8) - 3/8 \log_2(3/8) = 1.561$$

$$\text{infosep}(\text{weight}) = -2/8 \log_2(2/8) - 3/8 \log_2(3/8) - 3/8 \log_2(3/8) = 1.561$$

$$\text{infosep}(\text{lotion}) = -5/8 \log_2(5/8) - 3/8 \log_2(3/8) = 0.954$$

$$\text{GainRatio}(\text{hair}) = (0.954 - 0.5) / 1.406 = 0.321$$

$$\text{GainRatio}(\text{height}) = (0.954 - 0.689) / 1.561 = 0.1699$$

$$\text{GainRatio}(\text{weight}) = (0.954 - 0.939) / 1.561 = 0.0098$$

$$\text{GainRatio}(\text{lotion}) = (0.954 - 0.607) / 0.954 = 0.364$$

O atributo escolhido é "lotion".

- d) Calcule a razão de erro em cada uma das folhas da árvore que contém apenas o nó de decisão identificado na alínea anterior.

Aplicando o atributo "lotion":

- se lotion="no" então result="sunburned": 2 erros em 5 $\rightarrow (e+1)/(n+2) = 3/7 = 0.429$

- se lotion="yes" então result="none": 0 erros em 3 $\rightarrow (e+1)/(n+2) = 1/5 = 0.2$

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

- a) a) Num problema de pesquisa com um grande fator de ramificação, sabe-se que existem muitas soluções, todas elas a profundidades semelhantes, das quais sabemos a profundidade máxima. Queremos obter uma qualquer solução o mais rápido possível. Que método de pesquisa sistemática escolheria? Justifique.

Escolheria a pesquisa primeiro em profundidade com profundidade limitada, uma vez que permite obter uma solução a uma profundidade inferior ao limite dado, cujo valor máximo é sabido. A solução encontrada não é necessariamente a ótima, o que não é um requisito. Tem também poucos requisitos de memória, o que é importante dado o grande fator de ramificação do problema.

- b) Sabe-se de uma função heurística h que no máximo excederá em 10% o custo real. Defina, com base em h , a melhor função heurística admissível possível.

$$h' = h / 1.1 = 0.909 * h$$

- c) Um técnico de informática decidiu recuperar computadores antigos para criar uma *cloud*. A sua confiança na solução construída pode ser representada pelas seguintes regras: SE sw recente ENTÃO funciona (FC=0.8); SE computadores em_bom_estado ENTÃO funciona (FC=0.7); SE checkup ok ENTÃO computadores em_bom_estado (FC=0.75). O software (sw) utilizado é recente, e o check up parece ter corrido bem (FC=0.9). O que concluir, e com que fator de certeza?

$$R3: \text{computadores em_bom_estado} = 0.9 * 0.75 = 0.675$$

$$R2: \text{funciona com FC} = 0.675 * 0.7 = 0.4725$$

$$R1+R2: \text{funciona com FC} = 0.8 + 0.4725 * (1-0.8) = 0.8945$$

- d) Ao aplicar um algoritmo genético a um problema de maximização, obteve-se uma população com os cromossomas C1, C2, C3 e C4, com os valores 10, 15, 27 e 30, respetivamente. Determine a probabilidade de seleção de cada cromossoma.

$$10+15+27+30 = 82$$

$$p(C1) = 10/82 = 0.122$$

$$p(C2) = 15/82 = 0.183$$

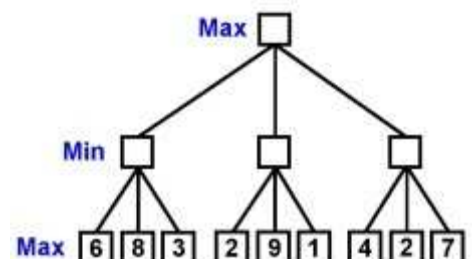
$$p(C3) = 27/82 = 0.329$$

$$p(C4) = 30/82 = 0.366$$

- e) Na figura ao lado, identifique os nós que são cortados pela aplicação de cortes alfa-beta ao algoritmo de pesquisa adversarial minimax.

- f) Um agente tem conhecimento completo do seu ambiente, aplicando este conhecimento num sistema baseado em regras. Indique se tais regras são causais ou de diagnóstico. Justifique.

Estas regras são causais, uma vez que o agente conhece todo o ambiente, daí sabendo quais as causas de cada perceção que tem do mesmo.



- g) Desenhe uma rede neuronal simples que consiga implementar a função paridade para 3 entradas binárias. O output da rede deve ser 1 se e só se houver um número ímpar de entradas a 1.

