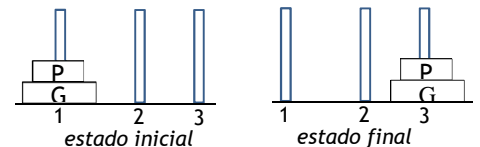


**Nota: Responder a cada questão (1, 2, 3 e 4) em folhas de exame separadas.**

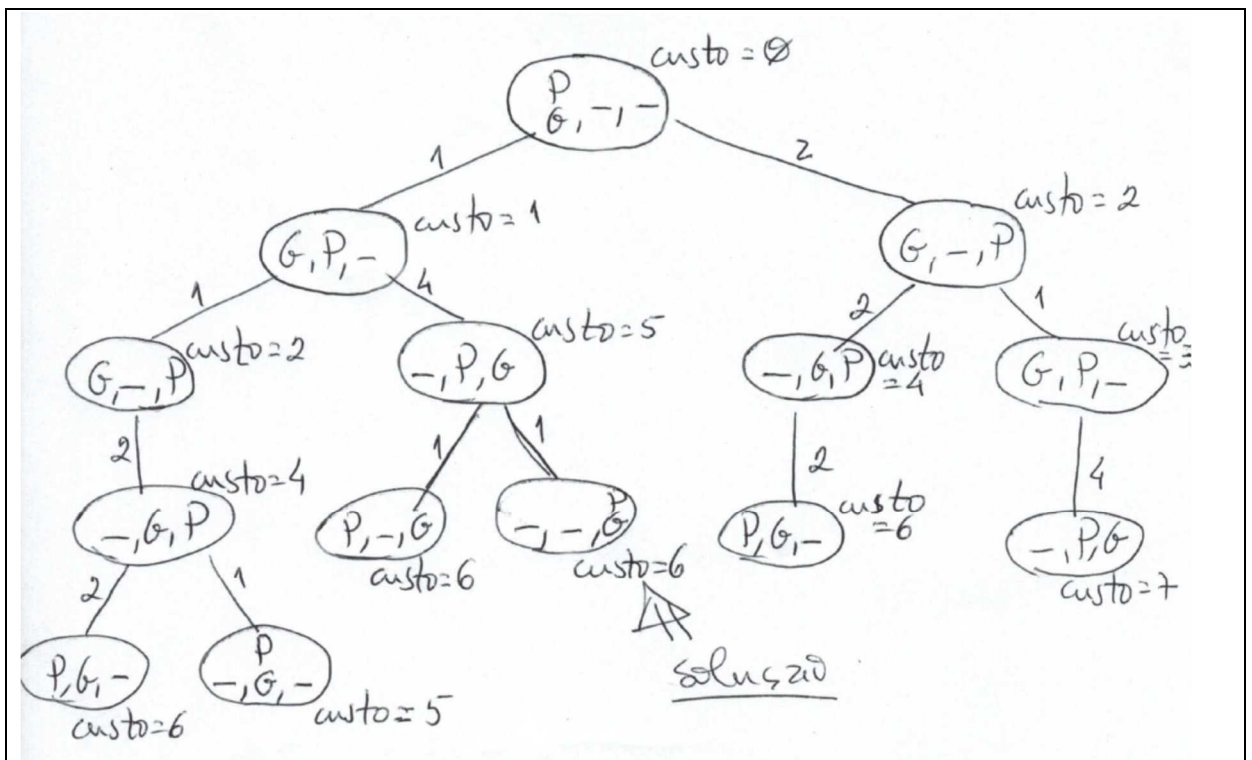
1. [4 valores] Torres de Hanói é um jogo que consiste em mover uma pilha de discos de tamanhos diferentes de uma haste inicial para uma haste final, com apoio de uma haste auxiliar. A figura seguinte apresenta o estado inicial e final do jogo para uma pilha de dois discos (versão a considerar).

Só pode ser movido um disco de cada vez (o disco que está no topo de uma pilha) e não pode ser colocado sobre um disco menor. Considere que o disco maior é designado por G e o disco menor por P. O disco G pesa 2 unidades e o disco P 1 unidade. O custo de mover um disco é igual ao seu peso a multiplicar pela distância que percorre (se haste adjacente, distância=1, senão distância=2). O objetivo do jogo é atingir o estado final com o menor custo.



- a) Como pode garantir que a estratégia de pesquisa **primeiro em profundidade**, aplicada a este cenário, encontre sempre uma solução? Justifique.

- b) Apresente a árvore de pesquisa gerada pela estratégia de **custo uniforme (BB)**, indicando os valores da função de avaliação em cada nó. Qual a solução encontrada?



- c) Na utilização da estratégia de pesquisa **A\***, pondera-se o uso das seguintes funções heurísticas:
- Número de discos na haste 3
  - Número de discos que não estão na haste 3
  - $\sum_i peso_i$ ,  $i$ : disco que não está na haste 3
  - $\sum_i peso_i * distancia_i$ ,  $i$ : disco que não está na haste 3
  - $\sum_i peso_i * distancia_i + penalizacao$ ,  $i$ : disco que não está na haste 3, penalizacao=1 se  $\exists i$  maior que disco no topo da haste 3, 0 senão

Quais destas heurísticas são admissíveis? Justifique.

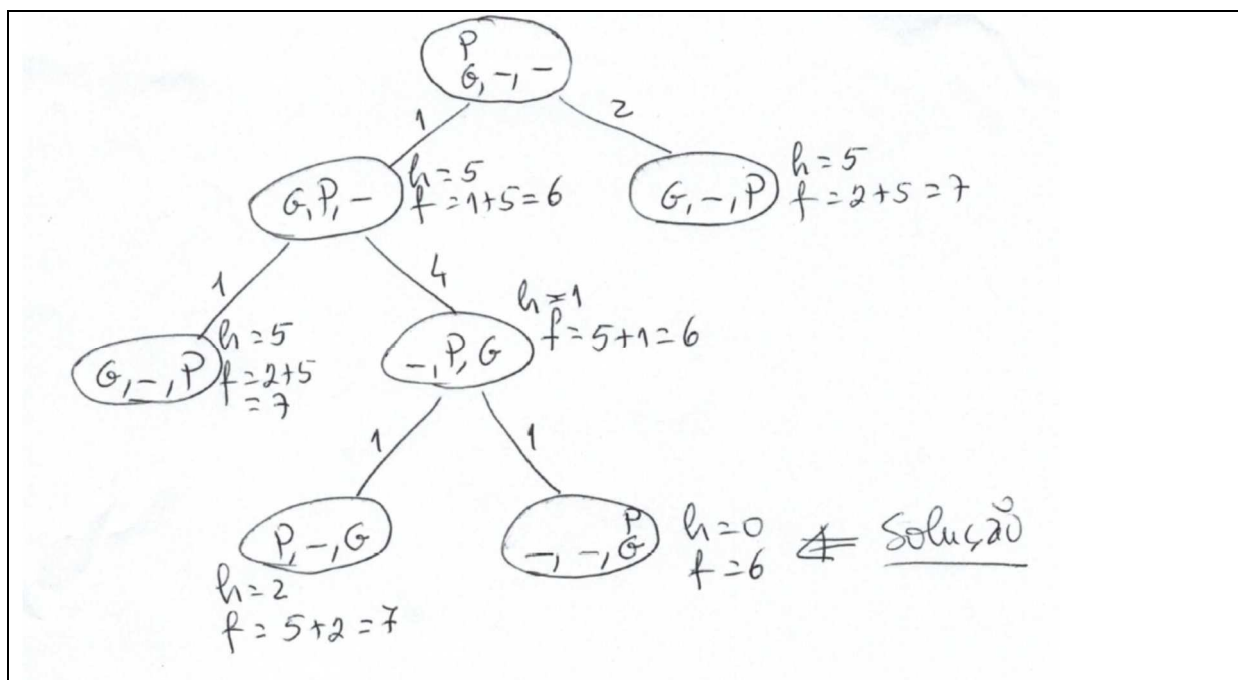
De entre as heurísticas admissíveis, qual considera a melhor? Justifique.

A função heurística i) não é admissível.

As funções heurísticas ii), iii), iv) e v) são admissíveis, pois nunca sobre-estimam o custo de chegar à solução.

A melhor heurística é a v), é a que se encontra mais perto do valor real.

- d) Apresente a árvore de pesquisa gerada pela estratégia A\* quando utiliza a heurística v. Indique os valores da função da avaliação em cada nó. Qual a solução encontrada?



2. [4 valores] O selecionador de uma determinada seleção nacional de futebol está preocupado com os resultados da sua equipa. Decidiu colecionar dados sobre os jogos disputados, de acordo com a tabela: se o adversário estava melhor ou pior classificado no ranking da FIFA, o tipo de jogo, a competição em causa e o estado do tempo que fazia. O selecionador tomou ainda nota do resultado do encontro.

Rank Adv	Tipo Jogo	Competição	Tempo	Resultado
melhor	fase final	européu	sol	vitória
melhor	qualificação	européu	sol	vitória
melhor	qualificação	mundial	chuva	derrota
melhor	fase final	mundial	sol	empate
melhor	amigável	nenhuma	chuva	empate
pior	fase final	européu	chuva	derrota
pior	qualificação	européu	chuva	vitória
pior	fase final	européu	sol	empate
pior	fase final	mundial	nublado	vitória
pior	amigável	nenhuma	nublado	empate

- a) Calcule a informação média para este conjunto de dados.

$$\text{info\_media} = -4/10 * \log_2(4/10) - 4/10 * \log_2(4/10) - 2/10 * \log_2(2/10) = 0.529 + 0.529 + 0.464 = 1.522$$

- b) Já se conhecem as entropias de cada atributo em relação ao resultado dos jogos: Rank Adversário (1,522), Tipo Jogo (1,036), Competição (1,161) e Tempo (1,2). Determine qual o atributo que maximiza a razão do ganho.

$$\text{gain}(\text{rank}) = 1.522 - 1.522 = 0$$

$$\text{split\_info}(\text{rank}) = -5/10 \cdot \log_2(5/10) - 5/10 \cdot \log_2(5/10) = 0.5 + 0.5 = 1$$

$$\text{gain\_ratio}(\text{rank}) = 0 / 1 = 0$$

$$\text{gain}(\text{tipo}) = 1.522 - 1.036 = 0.485$$

$$\text{split\_info}(\text{tipo}) = -2/10 \cdot \log_2(2/10) - 3/10 \cdot \log_2(3/10) - 5/10 \cdot \log_2(5/10) = 0.464 + 0.521 + 0.5 = 1.485$$

$$\text{gain\_ratio}(\text{tipo}) = 0.485 / 1.485 = 0.327$$

$$\text{gain}(\text{compet}) = 1.522 - 1.161 = 0.361$$

$$\text{split\_info}(\text{compet}) = -2/10 \cdot \log_2(2/10) - 5/10 \cdot \log_2(5/10) - 3/10 \cdot \log_2(3/10) = 0.464 + 0.5 + 0.521 = 1.485$$

$$\text{gain\_ratio}(\text{compet}) = 0.361 / 1.485 = 0.243$$

$$\text{gain}(\text{tempo}) = 1.522 - 1.2 = 0.322$$

$$\text{split\_info}(\text{tempo}) = -4/10 \cdot \log_2(4/10) - 2/10 \cdot \log_2(2/10) - 4/10 \cdot \log_2(4/10) = 0.529 + 0.464 + 0.529 = 1.522$$

$$\text{gain\_ratio}(\text{tempo}) = 0.322 / 1.522 = 0.212$$

O atributo que maximiza a razão do ganho é o "Tipo Jogo".

c) Construa a árvore de decisão, até uma profundidade máxima de 2 (raiz tem profundidade 0).

tipo=amigável => empate

tipo=fase\_final

$$\text{info\_media} = -2/5 \cdot \log_2(2/5) - 2/5 \cdot \log_2(2/5) - 1/5 \cdot \log_2(1/5) = 0.529 + 0.529 + 0.464 = 1.522$$

$$\text{entropia}(\text{rank}) = 2/5 \cdot [-1/2 \cdot \log_2(1/2) \cdot 2] + 3/5 \cdot [-1/3 \cdot \log_2(1/3) \cdot 3] = 0.4 + 0.951 = 1.351$$

$$\text{gain}(\text{rank}) = 1.522 - 1.351 = 0.171$$

$$\text{split\_info}(\text{rank}) = -2/5 \cdot \log_2(2/5) - 3/5 \cdot \log_2(3/5) = 0.529 + 0.442 = 0.971$$

$$\text{gain\_ratio}(\text{rank}) = 0.171 / 0.971 = 0.176$$

$$\text{entropia}(\text{compet}) = 3/5 \cdot [-1/3 \cdot \log_2(1/3) \cdot 3] + 2/5 \cdot [-1/2 \cdot \log_2(1/2) \cdot 2] = 0.951 + 0.4 = 1.351$$

$$\text{gain}(\text{compet}) = 1.522 - 1.351 = 0.171$$

$$\text{split\_info}(\text{compet}) = -3/5 \cdot \log_2(3/5) - 2/5 \cdot \log_2(2/5) = 0.442 + 0.529 = 0.971$$

$$\text{gain\_ratio}(\text{compet}) = 0.171 / 0.971 = 0.176$$

$$\begin{aligned} \text{entropia}(\text{tempo}) &= 1/5 \cdot [-1/1 \cdot \log_2(1/1)] + 1/5 \cdot [-1/1 \cdot \log_2(1/1)] + \\ &\quad 3/5 \cdot [-2/3 \cdot \log_2(2/3) - 1/3 \cdot \log_2(1/3)] = 0 + 0 + 0.551 = 0.551 \end{aligned}$$

$$\text{gain}(\text{tempo}) = 1.522 - 0.551 = 0.971$$

$$\text{split\_info}(\text{tempo}) = -1/5 \cdot \log_2(1/5) - 1/5 \cdot \log_2(1/5) - 3/5 \cdot \log_2(3/5) = 0.464 + 0.464 + 0.442 = 1.371$$

$$\text{gain\_ratio}(\text{tempo}) = 0.971 / 1.371 = 0.708$$

tipo=qualificação

$$\text{info\_media} = -2/3 * \log_2(2/3) - 1/3 * \log_2(1/3) = 0.390 + 0.528 = 0.918$$

$$\text{entropia}(\text{rank}) = 2/3 * [-1/2 * \log_2(1/2) * 2] + 1/3 * [-1/1 * \log_2(1/1)] = 0.667 + 0 = 0.667$$

$$\text{gain}(\text{rank}) = 0.918 - 0.667 = 0.252$$

$$\text{split\_info}(\text{rank}) = -2/3 * \log_2(2/3) - 1/3 * \log_2(1/3) = 0.390 + 0.528 = 0.918$$

$$\text{gain\_ratio}(\text{rank}) = 0.252 / 0.918 = \boxed{0.274}$$

$$\text{entropia}(\text{compet}) = 2/3 * [-2/2 * \log_2(2/2)] + 1/3 * [-1/1 * \log_2(1/1)] = 0 + 0 = 0$$

$$\text{gain}(\text{compet}) = 0.918 - 0 = 0.918$$

$$\text{split\_info}(\text{compet}) = -2/3 * \log_2(2/3) - 1/3 * \log_2(1/3) = 0.390 + 0.528 = 0.918$$

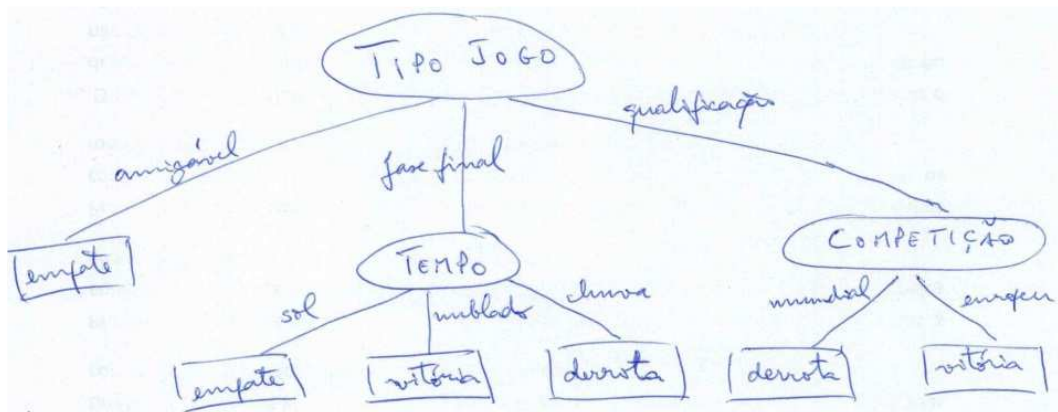
$$\text{gain\_ratio}(\text{compet}) = 0.918 / 0.918 = \boxed{1}$$

$$\text{entropia}(\text{tempo}) = 2/3 * [-1/2 * \log_2(1/2) * 2] + 1/3 * [-1/1 * \log_2(1/1)] = 0.667 + 0 = 0.667$$

$$\text{gain}(\text{tempo}) = 0.918 - 0.667 = 0.252$$

$$\text{split\_info}(\text{tempo}) = -2/3 * \log_2(2/3) - 1/3 * \log_2(1/3) = 0.390 + 0.528 = 0.918$$

$$\text{gain\_ratio}(\text{tempo}) = 0.252 / 0.918 = \boxed{0.274}$$



d) Calcule a razão do erro das folhas da árvore.

$$(e+1)/(n+2)$$

$$\text{tipo=amigável} \Rightarrow \text{empate: } (0+1)/(2+2) = 1/4 = 0.25$$

$$\text{tipo=fase\_final e tempo=sol} \Rightarrow \text{empate: } (1+1)/(3+2) = 2/5 = 0.4$$

$$\text{tipo=fase\_final e tempo=nublado} \Rightarrow \text{vitória: } (0+1)/(1+2) = 1/3$$

$$\text{tipo=fase\_final e tempo=chuva} \Rightarrow \text{derrota: } (0+1)/(1+2) = 1/3$$

$$\text{tipo=qualificação e compet=mundial} \Rightarrow \text{derrota: } (0+1)/(1+2) = 1/3$$

$$\text{tipo=qualificação e compet=europeu} \Rightarrow \text{vitória: } (0+1)/(2+2) = 1/4 = 0.25$$

3. [4 valores] Pretende-se minimizar o preço pago pela alocação de 3 lotes de terreno a três usos diferentes: Recreio (R), Habitação (H) e Lixeira (L). A área de Recreio deve estar localizada perto do lago e longe da auto-estrada. A área de Habitação não deve ser inclinada e deve estar longe da auto-estrada. A Lixeira não deve estar num terreno inclinado. Estão disponíveis 10 lotes de terreno, com as características indicadas na tabela.

Lote	Preço	Inclinação	Perto lago	Perto auto-estrada
L1	12K Eur	plano	não	sim
L2	13K Eur	plano	não	não
L3	9K Eur	plano	sim	sim
L4	16K Eur	plano	não	não
L5	17K Eur	plano	sim	não
L6	10K Eur	inclinado	não	não
L7	14K Eur	plano	sim	não
L8	8K Eur	inclinado	sim	não
L9	10K Eur	plano	sim	não
L10	20K Eur	plano	não	não

Aplica-se um **Algoritmo Genético** na resolução deste problema, sendo a população inicial constituída pelos 4 indivíduos:

(R=L5, H=L4, L=L3) ; (R=L7, H=L10, L=L5) ; (R=L5, H=L7, L=L4) ; (R=L9, H=L2, L=L1)

- a) Proponha uma estrutura para a representação do indivíduo, explicando. Exemplifique com a representação do 2º indivíduo da população inicial.

Representação de um indivíduo: 12 bits, cada 4 bits indicam o nº do lote atribuído a determinado uso. Cada conjunto de 4 bits refere-se a um uso: Recreio, Habitação, Lixeira

2º indivíduo da população inicial:

0 1 1 1 1 0 1 0 0 1 0 1

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

Uma possível função de adaptação = somatório dos custos de todos os lotes alocados.

- Se não é permitido que um indivíduo inclua terrenos não aptos a um uso, este valor é suficiente.

- Se é permitido que um indivíduo inclua terrenos não aptos a um uso, deve ser somado um valor de penalização (valor suficientemente elevado)

- 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): 0.87 / 0.38 / 0.58. Apresente o resultado deste processo de seleção. Explique.

I1: adaptação = 17 + 16 + 9 = 42

I2: adaptação = 14 + 20 + 17 = 51

I3: adaptação = 17 + 14 + 16 = 47

I4: adaptação = 10 + 13 + 12 = 35

Hipótese 1: converter o problema de minimização em problema de maximização

Subtrair o valor da adaptação ao valor máximo possível = 53 (soma dos 3 lotes mais caros)

I1: adaptaçãoM = 53 - 42 = 11

I2: adaptaçãoM = 53 - 51 = 2

I3: adaptaçãoM = 53 - 47 = 6

I4: adaptaçãoM = 53 - 35 = 18

Soma dos valores adaptação = 11+2+6+18 = 37

probabilidade seleção (I4) = 18/37 = 0.486 → intervalo ]0 ; 0.486]

probabilidade seleção (I1) = 11/37 = 0.297 → intervalo ]0.486 ; 0.783]

probabilidade seleção (I3) = 6/37 = 0.162 → intervalo ]0.783 ; 0.945]

probabilidade seleção (I2) = 2/37 = 0.054 → intervalo ]0.945 ; 1]

Resultado da seleção:

I4 (elitismo, é o mais apto) 1 0 0 1 0 0 1 0 0 0 0 1

I3 (valor aleatório=0.87)	0 1 0 1 0 1 1 1 0 1 0 0
I4 (valor aleatório=0.38)	1 0 0 1 0 0 1 0 0 0 0 1
I1 (valor aleatório=0.51)	0 1 0 1 0 1 0 0 0 0 1 1

- d) Calcule a 2ª geração da população, explicando. Sugira uma estratégia de cruzamento. A probabilidade de cruzamento é 75% e foram gerados os números aleatórios: 0.33 / 0.51 / 0.84. A probabilidade de mutação é 2% e só no 17º número aleatório surgiu um inferior a 0.02.

Estratégia de cruzamento: (por exemplo) cruzar os indivíduos num ponto de corte múltiplo de 4 (para não provocar alocações inválidas). Consideremos o ponto de corte entre os bits 8 e 9

1º filho:

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

2º filho:

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

(pela hipótese 1)

I4: mantém (elitismo)

I3: cruza (valor aleatório=0.33 < 0.75)

I4: cruza (valor aleatório=0.51 < 0.75)

I1: mantém (valor aleatório=0.84 > 0.75)

Após cruzamento:

I4 = 1 0 0 1 0 0 1 0 0 0 0 1

I34 = 0 1 0 1 0 1 1 1 0 0 0 1      mutação: 1ºbit mudaria, mas fica inválido, então não muda (é uma opção)

I43 = 1 0 0 1 0 0 1 0 0 1 0 0

I1 = 0 1 0 1 0 1 0 0 0 0 1 1

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

- a) No algoritmo de pesquisa por arrefecimento simulado, a aceitação de um novo estado depende da sua qualidade. Explique de que forma.

Um novo estado é sempre aceite quando a sua qualidade é melhor do que o estado atual. Caso contrário, a probabilidade da sua aceitação depende da diferença de qualidade entre esse novo estado e o estado atual: quanto menor for a diferença (isto é, quanto menos negativa), mais provável será que ele venha mesmo assim a ser aceite, de acordo com a fórmula  $e^{-\Delta E/\tau}$ , em que  $\tau$  é um parâmetro designado temperatura, que também influencia a aceitação do novo estado.

- b) Explique a importância do cálculo da plausibilidade no modelo Dempster-Shafer.

A Plausibilidade indica o valor máximo a que pode chegar a confiança numa determinada proposição, pelo eventual acrescento de mais evidências. A plausibilidade é dada por um menos a crença na negação da proposição em causa.

- c) No contexto da lógica dos conjuntos difusos, o que é uma função de pertença?

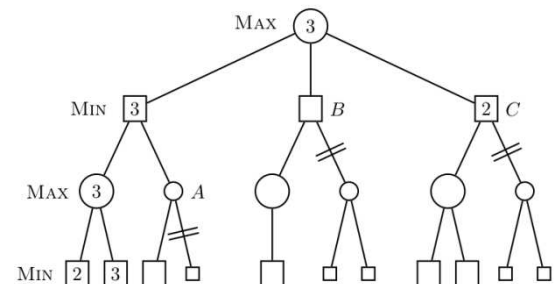
Uma função de pertença permite associar um valor a um conjunto difuso. Em vez de pertencer ou não a um conjunto, um valor tem um grau de pertença entre 0 e 1 dado por essa função, o que se traduz num grau de verdade de o valor em causa pertencer ou não ao conjunto difuso.

- d) Ao aplicar o algoritmo minimax, aplicaram-se os cortes alfa-beta indicados na figura (e só esses). Indique que gamas de valores podem ter os nós folha.

Folha "1":  $\geq 3$ .

Folha "2":  $\leq 3$ .

Folha "3" e "4": uma delas é 2, e a outra é  $\leq 2$ .



- e) Uma base de conhecimento indica que, para uma determinada equipa de futebol, um jogo de uma fase final de um campeonato do mundo resulta em vitória (FC=0,9). Por outro lado, se o jogo for de uma qualquer fase final e jogado com bom tempo, a vitória da equipa é provável (FC=0,7). A equipa vai jogar um jogo da fase final do campeonato do mundo, e as previsões meteorológicas indicam que vai estar bom tempo (FC=0,6). Aplique o modelo dos fatores de certeza para estimar a certeza na vitória.

R1: SE fase\_final E campeonato\_mundo ENTÃO vitória (0,9).

R2: SE fase\_final E bom\_tempo ENTÃO vitória (0,7).

F1: fase\_final.

F2: campeonato\_mundo.

F3: bom\_tempo (0,6).

Da aplicação de R1: vitória (0,9)

Da aplicação de R2: vitória ( $\min(1; 0,6) * 0,7 = 0,42$ )

Combinando R1 e R2:  $0,9 + 0,42 * (1 - 0,9) = 0,9 + 0,042 = 0,942$

A certeza na vitória é de 0,942.

- f) Num sistema de linguagem natural, é importante distinguir gramática e base de conhecimento. Qual a relação de cada um destes componentes com análise semântica e análise sintática?

A gramática permite definir o léxico e morfologia das palavras que se pretendem incluir em frases em linguagem natural. A partir da gramática podemos realizar análise sintática das frases. Já a base de conhecimento relaciona as palavras com o domínio sobre o qual expressamos determinadas frases (afirmativas ou interrogativas). É a partir da taxonomia (significado) e relacionamento das palavras definidas na base de conhecimento que se podem realizar diversas análises semânticas sobre o nexo das frases escritas em linguagem natural.

- g) Construiu-se uma rede neuronal com 30 neurónios de entrada, 1 camada escondida com 20 neurónios, e 2 neurónios na camada de saída. Os neurónios de cada camada ligam a todos os neurónios da camada seguinte. Em termos teóricos, quantos exemplos de treino são necessários para que a rede consiga generalizar?

Número de ligações independentes =  $30 \times 20 + 20 \times 2 = 600 + 40 = 640$

$640 < 2 \times NE \Leftrightarrow NE > 320$

Em termos teóricos, a rede só conseguirá generalizar com um mínimo de 320 exemplos de treino.