

Responda às questões apresentadas de forma clara e objectiva,
justificando sempre que necessário.

1.
 - 1.1. Relacione os principais processos internos de uma arquitectura de agente genérica. Indique, em pseudo-código, uma concretização desses processos, tal como implementado nas aulas práticas.
 - 1.2. Em que consiste o conceito de raciocínio prático e quais os principais conceitos com que se relaciona no contexto de uma arquitectura de agente. Indique, em pseudo-código, uma concretização desses conceitos, tal como implementado nas aulas práticas.
 - 1.3. Qual o papel do conceito de representação interna numa arquitectura de agente e como foi concretizada na prática? Indique o pseudo-código correspondente à respectiva utilização, bem como a sua função.
2. Considere uma fonte e dois vasos de água, um com capacidade de 3 litros, outro com capacidade de 4 litros, inicialmente vazios. É possível encher um vaso a partir da fonte ou de outro vaso, bem como despejar um vaso. Pretende-se que no vaso de 4 litros fiquem exactamente 2 litros de água.
 - 2.1. Em que consiste e qual a função da representação do problema no contexto de métodos de raciocínio automático baseados em procura em espaços de estados. Elabore uma representação do problema anterior.
 - 2.2. Tendo por base o método de *procura em largura*, elabore e apresente a árvore de procura, bem como as respectivas estruturas de dados auxiliares e a solução obtida. É possível a utilização do método de procura A* no contexto deste problema? Justifique.
 - 2.3. Qual o efeito da limitação da profundidade da procura nos métodos de *procura em largura* e de *procura em profundidade iterativa*. Indique, em pseudo-código, o modo de limitar a profundidade da procura.

3. Considere um ambiente caracterizado pela seguinte função de recompensa $R((x,y)) = \{ ((0,0), -20), ((1,0), -20), ((0,1), -20), ((1,1), -40), ((0,2), -20), ((1,2), 80) \}$ e no qual é possível a movimentação em quatro direcções: para a esquerda (\leftarrow), para a direita (\rightarrow), para cima (\uparrow), para baixo (\downarrow). A realização de uma acção tem uma probabilidade de sucesso de 50%, existindo uma probabilidade de 25% de desvio para cada uma das direcções perpendiculares ao movimento. A tentativa de movimento para uma posição não definida no ambiente deixa o sistema na mesma posição. Considere uma abordagem baseada em processos de decisão de Markov, com $\gamma = 0.5$ e $\Delta_{\max} = 15$.
- 3.1. Elabore um modelo do problema e calcule a política óptima.
 - 3.2. Relacione os principais conceitos envolvidos na resolução da alínea anterior e indique, em pseudo-código, o processo de cálculo utilizado.
 - 3.3. Num processo de decisão de Markov, qual a relação entre política óptima, limiar de convergência e factor de desconto temporal? Justifique.
4. Pretende-se implementar um agente capaz de aprender a movimentar-se no ambiente descrito na questão anterior, através de aprendizagem por reforço. Considere o conjunto de posições $P = \{(0,0), (1,0), (0,1), (1,1)\}$; a posição inicial $(0,1)$; a sequência de valores resultantes de um gerador de números aleatórios $R = [0.1, 0.2, 0.0, 0.4, 0.2, 0.7]$; a sequência de acções para efeitos de escolha aleatória $A = [\leftarrow, \uparrow, \rightarrow, \uparrow, \downarrow, \rightarrow]$; $\alpha = 0.5$; $\gamma = 0.5$, $\varepsilon = 0.3$. A tentativa de movimento para uma posição não definida em P deixa o sistema na mesma posição.
- 4.1. Tendo por base o método de aprendizagem *Q-Learning* e uma política de selecção de acção ε -greedy, simule a execução do algoritmo, indicando para cada iteração de aprendizagem os estados, acções e recompensas envolvidas, bem como os valores da função $Q(s,a)$ em cada iteração, para 5 iterações.
 - 4.2. Qual o papel do conceito de reforço no método de aprendizagem utilizado na alínea anterior? Indique, em pseudo-código, a função de reforço correspondente.
 - 4.3. Relacione o método de aprendizagem por reforço *Q-Learning* com o método de procura em espaço de estados A^* . Justifique.