Questões de Preparação para Exame

Inteligência Artificial - UBI 2025-26

BLOCO 1: Introdução à IA e Agentes Inteligentes

Questão 1.1

Explique as diferenças entre um agente que "age racionalmente" e um que "pensa racionalmente". Qual destas abordagens é mais adequada ao desenvolvimento científico da IA e porquê?

Resposta:

- **Pensar racionalmente**: Baseia-se na lógica e nos silogismos. O agente usa leis do pensamento para chegar a conclusões corretas. Problema: nem sempre é fácil transformar conhecimento informal em notação lógica, e problemas complexos podem esgotar recursos computacionais.
- **Agir racionalmente**: O agente escolhe a ação que maximiza o resultado esperado. É mais geral porque a inferência lógica é apenas um dos mecanismos para alcançar comportamento racional.
- A abordagem de **agir racionalmente** é mais adequada ao desenvolvimento científico porque usa uma definição clara e geral de racionalidade, é mais adaptável e não se limita apenas à inferência lógica.

Questão 1.2

Considere um aspirador robótico que limpa uma casa com 4 divisões. Classifique o ambiente deste problema de acordo com as seguintes propriedades, justificando cada resposta:

- Totalmente observável vs Parcialmente observável
- Determinístico vs Estocástico
- Episódico vs Sequencial
- Estático vs Dinâmico
- Discreto vs Contínuo
- Agente único vs Multi-agente

Resposta:

- **Parcialmente observável**: O aspirador pode não ter sensores em todas as divisões simultaneamente ou pode haver zonas não visíveis.
- **Estocástico**: O comportamento do aspirador e a sujidade podem variar (ex: pode não aspirar completamente, nova sujidade pode aparecer).
- **Sequencial**: As ações numa divisão afetam o estado global da casa e decisões futuras.
- **Dinâmico**: Nova sujidade pode aparecer enquanto o aspirador está a trabalhar.
- **Contínuo**: A posição do aspirador e quantidade de sujidade são valores contínuos.
- **Agente único**: Tipicamente apenas um aspirador (exceto se houver pessoas/animais a interferir).

BLOCO 2: Resolução de Problemas e Pesquisa

Questão 2.1

Para o problema do puzzle de 8 peças, proponha duas funções heurísticas admissíveis diferentes das apresentadas nas aulas (h_1 e h_2). Explique porque é que cada uma delas é admissível.

Resposta (exemplo):

- **h_3: Soma das distâncias de Manhattan dividida por 2**: É admissível porque cada movimento só pode colocar uma peça mais perto do seu destino, logo dividir por 2 nunca sobrestima.
- **h_4: Número de sequências lineares corretas**: Conta quantas vezes a próxima peça na sequência está na posição adjacente correta. É admissível porque nunca sobrestima o número de movimentos necessários para alcançar o objetivo.

Questão 2.2

Considere uma árvore de pesquisa com fator de ramificação b=4 e profundidade da solução d=6. Compare as complexidades temporal e espacial da Pesquisa Primeiro em Largura (PPL) com a Pesquisa Primeiro em Profundidade com Profundidade Iterativa (PPP-PI). Qual escolheria e porquê?

Resposta:

- **PPL**: Complexidade temporal O(b^d) = O(4^6) = O(4096); Complexidade espacial O(b^d) = O(4096)
- **PPP-PI**: Complexidade temporal O(b^d) = O(4096); Complexidade espacial O(bd) = O(24)
- **Escolha**: PPP-PI porque tem a mesma complexidade temporal que PPL mas complexidade espacial muito menor (24 vs 4096), sendo completa e ótima quando todas as ações têm o mesmo custo.

Questão 2.3

No contexto da pesquisa A*, explique: a) O que significa uma heurística ser "admissível"? b) Dê um exemplo de uma heurística não admissível para o problema de encontrar rotas num mapa. c) O que acontece se usarmos uma heurística não admissível na pesquisa A*?

Resposta: a) Uma heurística é admissível quando nunca sobrestima o custo real para atingir o objetivo, ou seja, h(n) ≤ custo_real(n, objetivo).

b) Exemplo não admissível: Distância euclidiana multiplicada por 2. Sobrestima sempre o custo real.

c) A pesquisa A* deixa de ser ótima, podendo devolver uma solução sub-ótima porque pode eliminar prematuramente caminhos que levam à solução ótima.

BLOCO 3: Pesquisa Adversarial (Jogos)

Questão 3.1

Considere a seguinte árvore de jogo com nodos de sorte (dado de 6 faces, todos os valores equiprováveis):

```
              MAX
         /     |     \
      MIN     MIN    MIN
      / \     / \    / \
   SORTE  4   3  SORTE  SORTE
    /|\        /|\      /|\
   2 8 4      1 5 7    6 3 9
```

Calcule o valor expectiminimax do nodo raiz. Mostre todos os cálculos.

Resposta:

- Nodos SORTE (da esquerda): (2+8+4)/3 = 14/3 = 4.67
- Nodo MIN (esquerdo): min(4.67, 4) = 4
- Nodos SORTE (do meio): (1+5+7)/3 = 13/3 = 4.33
- Nodo MIN (meio): min(3, 4.33) = 3
- Nodos SORTE (direita): (6+3+9)/3 = 18/3 = 6
- Nodo MIN (direito): min(6) = 6 (assumindo valor único)
- Nodo MAX (raiz): max(4, 3, 6) = **6**

Questão 3.2

Explique porque é que a função de avaliação num jogo com aleatoriedade (como o gamão) deve usar uma transformação linear positiva da utilidade, ao contrário dos jogos determinísticos onde podemos usar qualquer transformação monótona.

Resposta: Nos jogos com aleatoriedade, calculamos valores expectiminimax que dependem de médias ponderadas. Se usarmos transformações não lineares, a ordem das médias pode alterar-se, mudando a decisão ótima. Por exemplo, se transformarmos valores {1,4} e {2,3} com uma função quadrática: média de {1,4} = 2.5, mas média de {1,16} = 8.5; média de {2,3} = 2.5, mas média de {4,9} = 6.5. A ordem relativa inverte-se. Com transformações lineares positivas, a ordem das médias mantém-se.

Questão 3.3

Na poda alfa-beta, explique o significado dos parâmetros α e β. Em que situações é que podemos fazer poda e porquê é que isso não afeta o resultado final?

Resposta:

- **α**: Melhor valor (maior) que o MAX pode garantir até ao momento no caminho atual.
- **β**: Melhor valor (menor) que o MIN pode garantir até ao momento no caminho atual.
- **Poda**: Ocorre quando α ≥ β. Se estamos num nodo MIN e encontramos um valor ≤ α, o MAX nunca escolherá este ramo. Se estamos num nodo MAX e encontramos um valor ≥ β, o MIN nunca permitirá este ramo.
- **Resultado não afetado**: A poda apenas elimina ramos que comprovadamente não podem influenciar a decisão final. O valor minimax do nodo raiz mantém-se o mesmo.

BLOCO 4: Incerteza e Probabilidades

Questão 4.1

Um hospital pretende diagnosticar uma doença rara (D) que afeta 0.1% da população. Existe um teste (T) que deteta a doença com as seguintes características:

- Se a pessoa tem a doença, o teste é positivo em 99% dos casos: P(T|D) = 0.99
- Se a pessoa não tem a doença, o teste é positivo em 2% dos casos: P(T|¬D) = 0.02

a) Se uma pessoa faz o teste e este dá positivo, qual é a probabilidade de ela realmente ter a doença? Mostre todos os cálculos.

b) Explique porque é que o resultado pode ser contra-intuitivo.

Resposta:

a) Usando a Regra de Bayes:

- P(D) = 0.001
- P(¬D) = 0.999
- P(T|D) = 0.99
- P(T|¬D) = 0.02

P(T) = P(T|D)P(D) + P(T|¬D)P(¬D) = 0.99 × 0.001 + 0.02 × 0.999 = 0.00099 + 0.01998 = 0.02097

P(D|T) = P(T|D)P(D) / P(T) = (0.99 × 0.001) / 0.02097 = 0.00099 / 0.02097 = **0.047 ou 4.7%**

b) É contra-intuitivo porque apesar do teste ter 99% de precisão, a probabilidade de ter a doença dado teste positivo é apenas 4.7%. Isto deve-se à baixa prevalência da doença (0.1%) - há muitos mais falsos positivos do que verdadeiros positivos.

Questão 4.2

Explique a diferença entre probabilidade conjunta e probabilidade marginal, dando exemplos concretos usando as variáveis Cavidade e DorDentes das aulas.

Resposta:

- **Probabilidade conjunta P(A,B)**: Probabilidade de dois eventos ocorrerem simultaneamente. Exemplo: P(Cavidade, DorDentes) é a probabilidade de ter cavidade E dor de dentes ao mesmo tempo.

- **Probabilidade marginal P(A)**: Probabilidade de um único evento, obtida somando as probabilidades conjuntas sobre todas as possibilidades da outra variável. Exemplo: P(Cavidade) = P(Cavidade, DorDentes) + P(Cavidade, ¬DorDentes)

A marginalização "elimina" uma variável, somando sobre todos os seus valores possíveis.

BLOCO 5: Redes Bayesianas

Questão 5.1

Considere o seguinte cenário: Um estudante (E) pode estar cansado (C) ou não. Se estiver cansado, tem maior probabilidade de adormecer na aula (A). Se adormecer, o professor (P) pode reparar ou não, dependendo de quão atento está.

a) Desenhe a Rede Bayesiana para este problema. b) Indique quantos valores são necessários especificar para cada nodo (assuma variáveis booleanas). c) Escreva a expressão da probabilidade conjunta P(E,C,A,P) usando a estrutura da RB.

Resposta:

a) Rede Bayesiana:

```
 E
 |
 C
 |
 A
 |
 P
```

b) Valores necessários:

- E: 1 valor (P(E))
- C: 2 valores (P(C|E), P(C|¬E))
- A: 2 valores (P(A|C), P(A|¬C))
- P: 2 valores (P(P|A), P(P|¬A)) Total: 7 valores (vs 2^4-1 = 15 para distribuição conjunta completa)

c) P(E,C,A,P) = P(E) × P(C|E) × P(A|C) × P(P|A)

Questão 5.2

Dada a seguinte Rede Bayesiana sobre o desempenho de um website:

```
ServidorLento(S) → TempoCarregamento(T) ← TráfegoAlto(Tr)
                           ↓
                   UtilizadorInsatisfeito(U)
```

Com as seguintes probabilidades:

- P(S) = 0.1
- P(Tr) = 0.3
- P(T|S,Tr): V,V→0.95 | V,F→0.7 | F,V→0.6 | F,F→0.1

- P(U|T): V→0.9 | F→0.2

Calcule P(S|U), ou seja, a probabilidade do servidor estar lento dado que o utilizador está insatisfeito. Deixe indicados os cálculos que não conseguir efetuar.

Resposta:

Precisamos calcular P(S|U) = P(S,U) / P(U)

P(S,U) = Σ_{Tr,T} P(S) P(Tr) P(T|S,Tr) P(U|T)

Para S=V, U=V: = P(S=V) [P(Tr=V)P(T=V|S=V,Tr=V)P(U=V|T=V) + P(Tr=V)P(T=F|S=V,Tr=V)P(U=V|T=F) + P(Tr=F)P(T=V|S=V,Tr=F)P(U=V|T=V) + P(Tr=F)P(T=F|S=V,Tr=F)P(U=V|T=F)]

= 0.1 × [0.3×0.95×0.9 + 0.3×0.05×0.2 + 0.7×0.7×0.9 + 0.7×0.3×0.2] = 0.1 × [0.2565 + 0.003 + 0.441 + 0.042] = 0.1 × 0.7425 = 0.07425

P(¬S,U) calcula-se de forma similar com S=F: = 0.9 × [0.3×0.6×0.9 + 0.3×0.4×0.2 + 0.7×0.1×0.9 + 0.7×0.9×0.2] = 0.9 × [0.162 + 0.024 + 0.063 + 0.126] = 0.9 × 0.375 = 0.3375

P(U) = P(S,U) + P(¬S,U) = 0.07425 + 0.3375 = 0.41175

P(S|U) = 0.07425 / 0.41175 = 0.180 ou 18%

Questão 5.3

Explique o que é o "lençol de Markov" de uma variável numa Rede Bayesiana e qual a sua importância para a Amostragem de Gibbs.

Resposta: O **lençol de Markov** de uma variável X é composto por:

- Os seus pais diretos
- Os seus filhos diretos
- Os outros pais dos seus filhos (co-pais)

Importância: A distribuição de probabilidade de uma variável depende APENAS das variáveis no seu lençol de Markov. Na Amostragem de Gibbs, quando amostramos uma variável, só precisamos considerar os valores das variáveis no seu lençol de Markov, tornando o cálculo muito mais eficiente. Isso permite calcular P(X|resto da rede) = P(X|LM(X)).

BLOCO 6: Raciocínio Temporal

Questão 6.1

No problema do guarda-chuva, assumimos um processo de Markov de primeira ordem. Explique: a) O que significa ser um processo de Markov de primeira ordem? b) Quais seriam as vantagens e desvantagens de usar um processo de segunda ordem? c) Como poderia melhorar o modelo sem aumentar a ordem do processo de Markov?

Resposta:

a) **Processo de Markov de primeira ordem**: O estado atual depende apenas do estado imediatamente anterior: P(Xt|X0:t-1) = P(Xt|Xt-1). O passado mais distante não influencia diretamente o presente se conhecermos o estado anterior.

b) **Segunda ordem** (P(Xt|Xt-2, Xt-1)):

- Vantagens: Modelo mais realista, captura padrões temporais mais complexos (ex: tendências de 2 dias).
- Desvantagens: Mais parâmetros para estimar (ex: tabela com 4 entradas em vez de 2 para variáveis booleanas), maior custo computacional, risco de overfitting com poucos dados.

c) **Melhorias sem aumentar ordem**:

- Adicionar mais variáveis de estado (ex: estação do ano, pressão atmosférica)
- Usar mais sensores (ex: sensor de humidade, barómetro)
- Melhorar as probabilidades condicionais com mais dados históricos

---

Questão 6.2

Explique a diferença entre filtragem, previsão e suavização no contexto do raciocínio temporal. Dê um exemplo prático de quando cada uma seria utilizada.

Resposta:

- **Filtragem**: P(Xt|e1:t) - Estima o estado atual dadas todas as observações até agora.
  - Exemplo: Um robot navegando usa filtragem para saber onde está AGORA com base em todos os sensores até ao momento.
- **Previsão**: P(Xt+k|e1:t), k>0 - Estima um estado futuro.
  - Exemplo: Previsão meteorológica para daqui a 3 dias com base em dados históricos até hoje.
- **Suavização**: P(Xt-k|e1:t), 0<k≤t - Estima um estado passado usando também observações futuras.
  - Exemplo: Análise forense de um crime - estimamos o que aconteceu ontem usando também evidências encontradas hoje. Dá melhores estimativas que a filtragem porque usa mais informação.

Questão 6.3

Na filtragem recursiva, derivámos a seguinte equação:

$$P(X_t|e_{1:t}) = \alpha P(e_t|X_t) \Sigma_{x_{t-1}} P(X_t|x_{t-1}) P(x_{t-1}|e_{1:t-1})$$

Explique: a) O que representa cada termo desta equação? b) Porque é que esta formulação recursiva é vantajosa? c) O que aconteceria se não usássemos esta abordagem recursiva?

Resposta:

a) **Termos**:

- α: Constante de normalização para garantir que as probabilidades somam 1
- $P(e_t|X_t)$: Modelo de sensor - probabilidade da observação atual dado o estado atual
- $P(X_t|x_{t-1})$: Modelo de transição - como o estado evolui
- $P(x_{t-1}|e_{1:t-1})$: Termo recursivo - nossa crença sobre o estado anterior (resultado da iteração anterior)
- Σ: Marginalização sobre todos os possíveis estados anteriores

b) **Vantagens da recursão**:

- Custo computacional constante a cada passo (não cresce com t)
- Não é necessário guardar todo o histórico de observações
- Apenas precisamos do resultado da iteração anterior
- Eficiente em termos de memória e tempo

c) **Sem recursão**: Teríamos que recalcular usando todas as observações desde o início a cada passo. O custo cresceria linearmente com o tempo ($O(t)$ em vez de $O(1)$), tornando inviável para sistemas que operam durante longos períodos.

QUESTÕES INTEGRADORAS

Questão I.1

Compare a utilidade da lógica proposicional com as redes Bayesianas para representar conhecimento incerto. Quando devemos usar cada uma? Ilustre com um exemplo concreto.

Resposta:

Lógica Proposicional:

- Boa para conhecimento certo e regras determinísticas
- Não lida bem com incerteza
- Exemplo: "Se está a chover E não tenho guarda-chuva, então fico molhado" - regra determinística

Redes Bayesianas:

- Especializadas em conhecimento incerto
- Capturam dependências probabilísticas
- Permitem inferência com informação incompleta
- Exemplo: "Se está nublado, há 80% de probabilidade de chover" - modelação de incerteza

Quando usar:

- Lógica: Sistemas com regras claras e determinísticas (ex: sistemas de controlo simples, verificação formal)
- RB: Diagnóstico médico, previsão meteorológica, sistemas de recomendação - onde há incerteza inerente e múltiplos fatores probabilísticos

---

Questão I.2

Um agente inteligente está a jogar um jogo estocástico (com dados) e determinístico (xadrez) simultaneamente - em cada turno joga uma jogada de cada jogo. Como combinaria as técnicas de pesquisa adversarial estudadas para criar uma estratégia integrada? Que desafios teria que resolver?

Resposta:

Abordagem híbrida:

1. **Para o jogo estocástico**: Usar expectiminimax para calcular valores esperados considerando probabilidades dos dados
2. **Para xadrez**: Usar minimax com poda alfa-beta
3. **Integração**: Combinar as utilidades dos dois jogos numa função de utilidade global, possivelmente ponderada

Desafios:

- **Escalas diferentes**: Normalizar utilidades dos dois jogos
- **Trade-offs**: Decidir quando sacrificar posição num jogo para ganhar no outro
- **Custo computacional**: Dois espaços de pesquisa simultaneamente - necessidade de profundidades de pesquisa limitadas
- **Correlação**: Pode haver interdependências estratégicas entre os jogos
- **Horizonte temporal**: Os jogos podem ter durações diferentes

Solução:

- Função de avaliação: $U_{total} = w_1 \times U_{estocástico} + w_2 \times U_{xadrez}$
- Pesquisa com profundidade limitada em ambos
- Ajuste dinâmico dos pesos w_1 e w_2 conforme o estado dos jogos

---

Questão I.3

Discuta as limitações dos modelos de Markov de primeira ordem no raciocínio temporal. Dê três exemplos de situações reais onde estas limitações seriam problemáticas e sugira possíveis soluções.

Resposta:

Limitações:

- Assume que o passado só influencia o presente através do estado imediatamente anterior
- Não captura tendências ou padrões de longo prazo
- Pode perder informação importante de eventos distantes

Exemplos problemáticos:

1. **Mercado de ações**:
    - Problema: Tendências de longo prazo (bull/bear markets) não são capturadas
    - Solução: Usar ordem superior ou adicionar variável de estado "tendência de longo prazo"
2. **Saúde do paciente**:
    - Problema: Tratamento de 3 meses atrás pode afetar estado atual, mas não afeta o estado de ontem
    - Solução: Adicionar variáveis de estado que representem histórico de tratamentos
3. **Clima/meteorologia**:
    - Problema: Padrões sazonais e ciclos não são bem modelados
    - Solução: Adicionar variável de estado "estação do ano" ou usar modelos com período sazonal

Soluções gerais:

- Aumentar ordem do modelo (mas aumenta complexidade)
- Adicionar variáveis de estado que resumam histórico relevante
- Usar modelos híbridos que combinam diferentes horizontes temporais

---

FIM DO DOCUMENTO

Total de questões: 23 questões abrangendo todos os tópicos principais