

Iniciado em sábado, 21 Nov 2020, 05:22

Estado Finalizada

Concluída em sábado, 21 Nov 2020, 05:26

**Tempo
empregado** 3 minutos 24 segundos

Avaliar 5,00 de um máximo de 5,00(100%)

Questão **1**

Correto

Atingiu 1,00 de
1,00

Em relação a determinação de uma função de recompensa, julgue as afirmativas a seguir.

I. Na aprendizagem por reforço um agente aprendiz tem como objetivo maximizar o total de recompensa recebida ao longo do processo.

CERTO



II. Para atingir o objetivo, o agente não necessita mensurar quão bom é escolher uma determinada ação a partir do estado corrente por meio de um valor numérico, isso é feito pela função de recompensa.

ERRADO



III. Na aprendizagem por reforço o objetivo de um agente aprendiz é minimizar o total de recompensa recebida ao longo do processo.

ERRADO



IV. Para atingir o objetivo o agente necessita mensurar quão bom é escolher uma determinada ação a partir do estado corrente por meio de um valor numérico, isso é feito pela função de recompensa.

CERTO



Sua resposta está correta.









Questão **2**

Correto

Atingiu 1,00 de 1,00

Ao longo da disciplina, estudamos detalhes do algoritmo *Q-learning*. Faça a correta associação entre as proposições abaixo e as políticas correspondente.

1. Tem como objetivo estabelecer o comportamento do agente aprendiz para que ele alterne adequadamente entre o uso do conhecimento já adquirido e a aquisição de novo conhecimento.
2. Escolhe a ação que tem o maior valor esperado, com probabilidade definida por $(1 - \epsilon)$, e de ação aleatória, com probabilidade ϵ .
3. Permite escolher a ação que tem o maior valor esperado, com probabilidade definida por $(1 - \epsilon)$, e ação aleatória, com probabilidade. Além disso, o valor sofre um decaimento exponencial.
4. A escolha de ações é feita baseada em uma técnica denominada Comparação de Reforço.

-   Política baseada na contagem de visitas
-   Política ϵ - gulosa adaptativa
-   Política ϵ - gulosa
-   Políticas de seleção de ações para o algoritmo Q-Learning

Sua resposta está correta.


Questão **3**

Correto

Atingiu 1,00 de 1,00

Q-learning consiste em um tipo específico de aprendizagem por reforço que atribui valores a pares de estado-ação.

Escolha uma opção:

- ☒ Verdadeiro 
- ☐ Falso

Questão **4**

Correto

Atingiu 1,00 de 1,00

Sobre os tipos de aprendizagem Q, assinale a alternativa incorreta.

Escolha uma:

- ☐ a. A aprendizagem Q baseada em política tenta aprender, a função Q para alguma política projetada.
- ☐ b. A aprendizagem Q é um método de política-off.
- ☒ c. A aprendizagem Q é um método de política-off. A função valor ótima não é estimada, independentemente da política atual que está sendo utilizada para gerar as trajetórias amostradas. ✓ A letra C está errada já que a aprendizagem Q é um método de política-off, o que significa que a função valor ótima é estimada, independentemente da política atual (exploração) que está sendo utilizada para gerar as trajetórias amostradas.
- ☐ d. A função valor ótima da aprendizagem Q é estimada, independentemente da política atual que está sendo utilizada para gerar as trajetórias amostradas
- ☐ e. Na aprendizagem Q baseada em política, política pode ser ou não a política que de fato está sendo seguida durante o treinamento.

Sua resposta está correta.

Questão **5**

Correto

Atingiu 1,00 de 1,00

A ideia de experimentar mais de uma política de seleção de ações para o *Q-learning*, tem como meta verificar qual dessas políticas é mais adequada para ser utilizada na implementação dos métodos híbridos propostos.

Escolha uma opção:

- ☒ Verdadeiro ✓
- ☐ Falso