

Iniciado em	quinta, 5 Nov 2020, 10:49
Estado	Finalizada
Concluída em	quinta, 5 Nov 2020, 10:54
Tempo empregado	5 minutos 52 segundos
Avaliar	10,00 de um máximo de 10,00(100%)

Questão 1

Correto

Atingiu 2,00 de 2,00

Em relação aos métodos que compreendem a aprendizagem passiva, julgue as afirmativas a seguir.

I. A ideia da programação dinâmica adaptativa é aprender o modelo de transição e a função de reforço empiricamente. Para isso, ela usa a iteração de valor simplificada para obter as utilidades. CERTO ✓

II. Os algoritmos de aprendizado por diferença temporal aprendem novas estimativas do valor com base em outras estimativas, não exigindo um modelo exato do sistema. CERTO ✓

III. Uma vez que calculamos a utilidade de um estado como recompensa, a utilidade de cada estado pode ser escrita estritamente em termos das utilidades de seus vizinhos imediatos. Essa afirmação diz respeito aos métodos de estimativa de utilidade direta. CERTO ✓

IV. No método estimativa de utilidade direta um agente PDA obtém proveito das limitações entre as utilidades de estados compreendendo o modelo de transição que os conecta e soluciona o processo de decisão de Markov correspondente. ERRADO ✓

Sua resposta está correta.

A proposição IV está incorreta. Um agente PDA obtém proveito das limitações entre as utilidades de estados compreendendo o modelo de transição que os conecta e soluciona o processo de decisão de Markov correspondente, fazendo uso de um método de **programação dinâmica**.

As demais proposições estão corretas.

Questão **2**

Correto

Atingiu 2,00 de 2,00

A aprendizagem por reforço passiva faz uso de um conceito baseado em estados em um ambiente completamente observável. O ambiente passa de um estado para outro sem intervenção do agente, em que a cada transição o agente apenas percebe o estado e recebe reforço, atualizando sua representação do valor desse estado.

Escolha uma opção:

- ☒ Verdadeiro ✓
- ☐ Falso

Questão **3**

Correto

Atingiu 2,00 de 2,00

Em relação a aprendizagem por reforço passiva o agente tenta localizar uma boa política de ação. A cada transição, o agente compreende o estado e obtém o reforço, atualizando sua interpretação do valor nesse estado.

Escolha uma opção:

- ☐ Verdadeiro
- ☒ Falso ✓

Questão 4

Correto

Atingiu 2,00 de 2,00

Sobre aprendizagem por reforço passiva, julgue os itens a seguir escolhendo a única alternativa correta.

Escolha uma:

- ☐ a. Um agente deve escolher quais ações deve executar, analogamente a aprendizagem por reforço passiva.
- ☒ b. O agente precisa aprender um modelo completo com probabilidades de resultados para todas as ações, em vez de aprender apenas o modelo para a política fixa. ✓ Na aprendizagem ativa, o agente precisa aprender um modelo completo com probabilidades de resultados para todas as ações, em vez de aprender apenas o modelo para a política fixa. Um agente deve escolher quais ações deve executar, de modo diferente da aprendizagem passiva. Um agente deve assumir um compromisso entre exploração (aproveitamento) para maximizar a sua recompensa (que reflete em suas estimativas de utilidades atuais) e exploração, a fim de maximizar seu bem-estar a longo prazo.
- ☐ c. Na aprendizagem por reforço ativa, um agente deve assumir um compromisso entre exploração e estimativa de utilidade direta.
- ☐ d. O agente não precisa aprender um modelo completo com probabilidades de resultados para todas as ações, em vez de aprender apenas o modelo para a política fixa.
- ☐ e. Na aprendizagem por reforço ativa, um agente não deve assumir um compromisso entre exploração e estimativa de utilidade direta.

Sua resposta está correta.

Questão 5

Correto

Atingiu 2,00 de 2,00

Sobre aprendizagem por reforço passiva e aprendizagem por reforço ativa, faça as associações corretas de acordo com os conceitos vistos nesta Unidade:

1. Aprendizagem por reforço passiva.

2. Aprendizagem por reforço ativa.

- ✓ Compreendem os métodos: estimativa de utilidade direta, programação dinâmica adaptativa e diferença temporal.
- ✓ As duas principais estratégias encontradas na literatura para aprendizagem por reforço utilizando DT são as seguintes: crítico heurístico adaptativo e *Q-learning*.
- ✓ Exploração e aprendizagem de uma função de ação.
- ✓ Um agente guloso raramente converge para a política ótima correspondente ao ambiente.

Sua resposta está correta.

Aprofunde mais sobre essa questão assistindo as aulas da Unidade III.