Luis Felipe Schalcher de Oliveira - trabalho de aprendizado por reforço - Introdução à robótica

Agente: robô
Ambiente: obstáculo, destino e espaço disponível para a locomoção
Ações: andar nas 8 possíveis direções e ficar parado
Variáveis de estado: posição do robô e posição do obstáculo
Algoritmo usado: Q-learning
Constante alpha: 0.05
Constante gamma: 0.5

Como o programa funciona:

- O programa faz o uso de bibliotecas de Python para facilitar a execução do código, que são: numpy, random e matplotlib. Depois das importações, o programa inicializa as variáveis e constantes que serão usadas e a matriz Q, totalmente zerada, a princípio. Após isso, algumas funções são definidas:
- recompensa: define quais serão as recompensas e punições, dependendo da ação escolhida
- escolher_acao: define a ação a ser escolhida de acordo com a política epsilon-greedy
- mover_obstaculo: faz com que o obstáculo se mova de acordo com a trajetória proposta
- mover_robo: define 9 ações, enumeradas de 0 a 8, correspondentes às possibilidades de movimentação do robô
- Após a definição das funções, o treinamento é iniciado e, até que o número de episódios escolhidos seja completo, o robô e o obstáculo se moverão conforme o proposto, até que o robô chegue no destino ou colida. A matriz Q é atualizada de acordo com a função da política de Q-learning e as recompensas vão sendo aplicadas conforme o desenrolar dos episódios. A cada episódio, a recompensa final é adicionada a uma lista, que será usada para a impressão do gráfico recompensa/episódio posteriormente.
- Depois que o treinamento é concluído, o gráfico mencionado é exibido, e o robô realiza uma trajetória para chegar no destino que, após o número de episódios escolhido (100 mil), em todos os casos de teste, foi feita no número mínimo de passos. As coordenadas que o robô e o obstáculo percorreram durante a execução do programa são exibidas no terminal e uma função que anima a trajetória do robô e do obstáculo é executada, para vê-la, basta fechar a janela com o gráfico, que se sobrepõe a ela.