

# Trabalho Prático I - Problema de Transporte de Objeto

Disciplina: Reinforcement Learning  
Curso: Especialização em Ciência de Dados  
Professor: Ricardo Grunitzki

19/06/2021

## Descrição

Considere o problema de mundo em grade ilustrado a seguir:

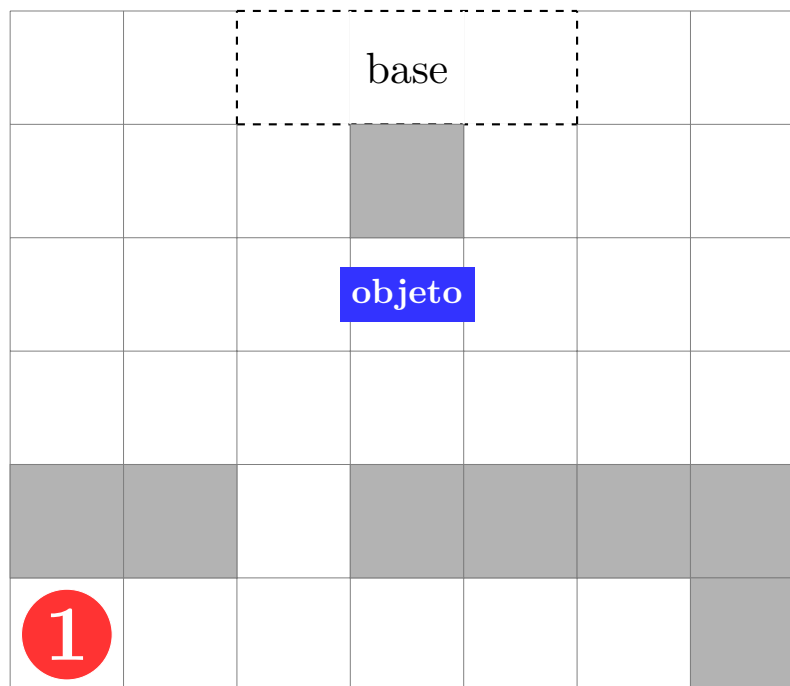


Figure 1: Cenário do Problema de Transporte de Objeto

Neste cenário, um agente (em vermelho) deve percorrer a grade 7x6, encontrar o objeto e transportá-lo até na base. Essa tarefa deve ser executada na menor quantidade de passos de tempo possível. O agente não possui nenhum conhecimento prévio sobre o ambiente, o qual possui paredes, representadas pelas células pintadas de cinza, as quais ele não transpor. O agente também não possui conhecimento prévio sobre a localização do objeto.

A localização inicial do agente, disposição das paredes e objeto são sempre fixas, conforme indicado na ilustração. A cada passo de tempo, o agente pode executar os seguintes movimentos na grade:

- mover para cima;

- mover para baixo;
- mover para esquerda;
- mover para a direita;
- permanecer na mesma célula.

Este cenário apresenta algumas restrições de movimentação:

- O agente pode realizar apenas uma movimentação por passo de tempo.
- Se o agente escolher se mover para uma célula que não está vazia, seja por conta de uma parede ou objeto, ele não se move, i.e., permanece na mesma célula.
- Qualquer tentativa de locomoção para além da grade, resultará na não movimentação do agente.
- O objeto só pode ser agarrado pela sua esquerda ou direita.
- Quando o agente é posicionado à direita ou esquerda do objeto, o objeto é agarrado automaticamente.
- Uma vez agarrado o objeto, o agente não pode solta-lo;
- O agente, quando agarrado ao objeto, só consegue se mover para uma nova célula desde que não haja nenhuma restrição de movimentação para o agente e objeto.

O episódio é concluído automaticamente quando o objeto entra na base ou se atingir um número máximo de passos de tempo sem resolver a tarefa. Em ambos os casos, um novo episódio é reiniciado, com o agente e objeto situados conforme Figura 1.

## Atividade

Implemente uma solução via reinforcement learning para o problema de transporte de objeto e apresente um relatório endereçando os seguintes aspectos da solução:

### 1. Modelagem do MDP:

- Apresente a modelagem de estados considerada, bem como a quantidade de estados presentes no MDP. Inclua na contagem os estados não-válidos;
- Apresente a modelagem das ações que o agente pode executar;
- Apresente a modelagem da função de recompensa, com as situações em que o agente é recompensado bem como a magnitude da recompensa. Justifique as suas escolhas.

### 2. Configuração dos Experimentos

- Apresente os valores de taxa de aprendizagem (alfa) e fator de desconto (gamma) do algoritmo de aprendizagem Q-Learning;
- Apresente as configurações do horizonte de aprendizagem, que é representado pela quantidade máxima de passos de tempo por episódios, quantidade máxima de episódios, e política de exploração ao longo do tempo;

### 3. Resultados Experimentais

- (a) Apresente a curva de convergência, representada pela quantidade de passos (timesteps) necessários para resolver a tarefa ao longo do tempo (episódios).
- (b) Apresente o tempo de processamento necessário para resolver o problema.

## Entrega

A atividade deverá ser entregue até **26/06/2021**, às 07:59, no AVA3. Você deve entregar um arquivo zip contendo:

- O código-fonte da solução;
- Instruções para a execução do código; e
- Um relatório em PDF contendo as respostas dos questionamentos.

Fique à vontade para utilizar a linguagem de programação e estrutura de dados de sua preferência.