

Tema:

Aprendizagem por reforço aplicada ao problema do pêndulo invertido

Neste trabalho foi estudada a aplicação de métodos de aprendizagem por reforço ao problema do pêndulo invertido, um problema de controle bastante discutido cujo objetivo é de equilibrar um pêndulo, acoplado a uma plataforma móvel, em sua posição instável. A aprendizagem por reforço é um ramo da inteligência artificial que foca no estudo de agentes que, sem nenhum conhecimento prévio do ambiente com o qual interagem, devem aprender a maximizar a recompensa obtida.

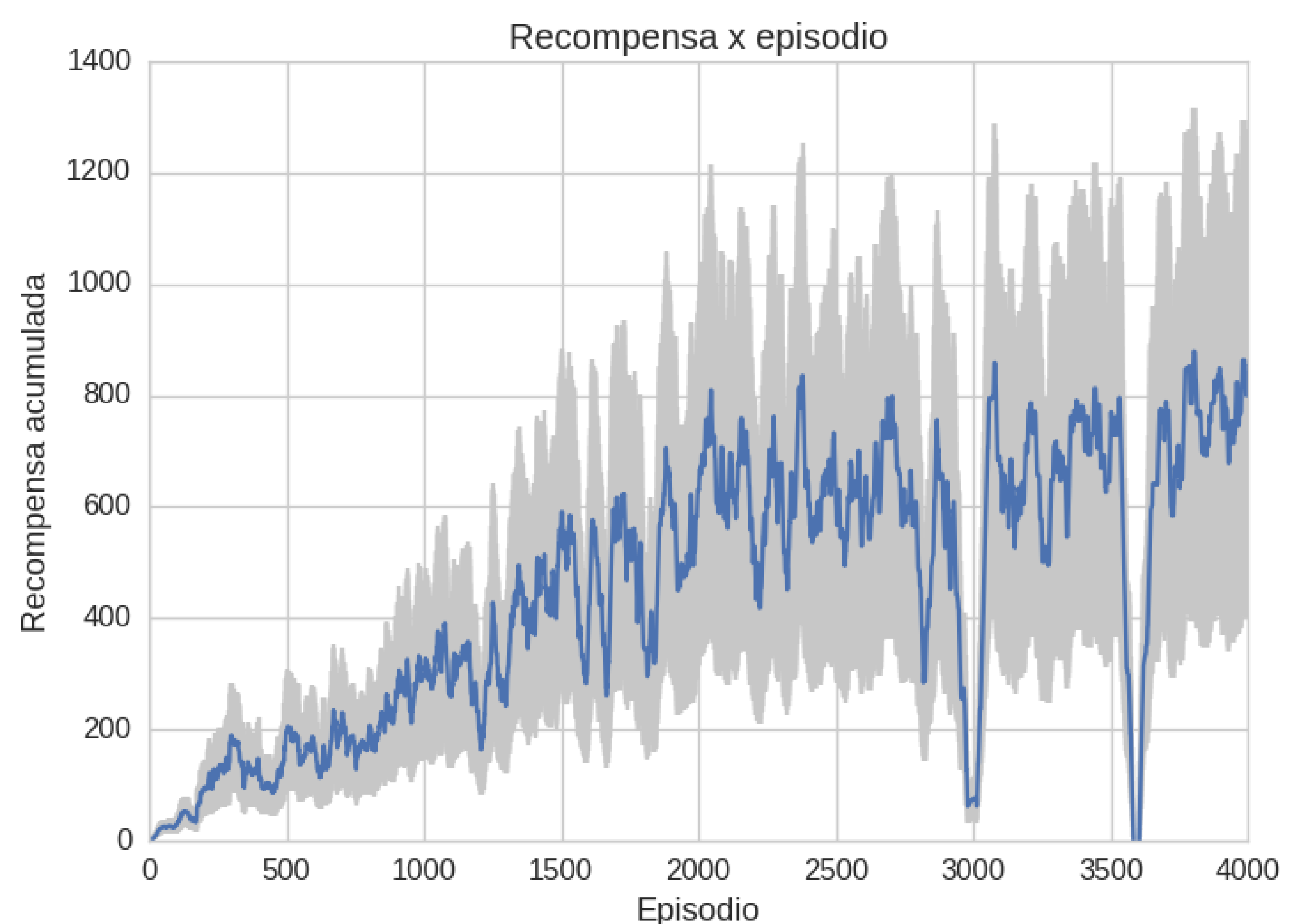
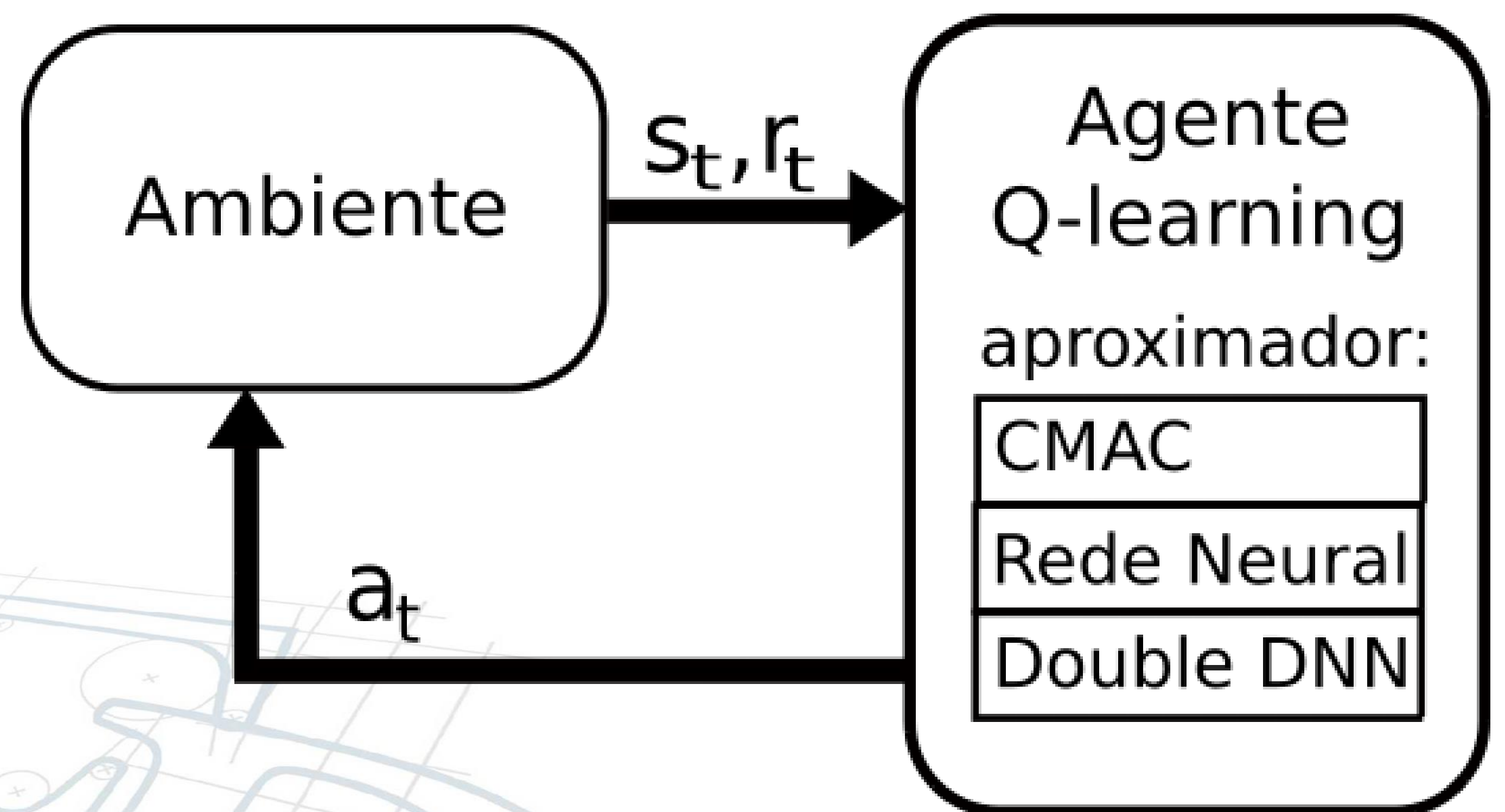
A cada instante de tempo o ambiente fornece ao agente o estado atual no qual este se encontra e uma recompensa associada a um par estado x ação. O agente deve então decidir uma ação e aplicá-la ao ambiente.

Focamos principalmente no algoritmo de programação dinâmica Q-learning, que tenta aproximar a recompensa esperada de um par estado x ação a partir de suas experiências passadas. O agente então toma decisões que maximizam a recompensa esperada em um dado estado.

Em sua versão original o algoritmo utiliza uma tabela para aproximar a recompensa esperada, no entanto em problemas reais o número de estados possíveis pode se tornar muito grande o que torna inviável uma representação tabular destes valores. Utiliza-se então aproximadores desta função, como por exemplo o CMAC (Cerebellar Model Articulation Controller) ou Redes Neurais associados a este algoritmo.

Desenvolveu-se então um simulador do sistema do pêndulo invertido com uma interface gráfica na linguagem Python utilizando a plataforma OpenAI-gym, foram aplicados diversos algoritmos de aprendizagem por reforço ao sistema estudado e obtivemos bons resultados utilizando redes neurais como um aproximador do algoritmo Q-learning.

Este método foi então adaptado de modo a trabalhar em conjunto com um sistema físico do pêndulo invertido disponibilizado pelo Laboratório de Automação e Controle da USP e foi possível assim observar diversos obstáculos que surgem nesta migração ao sistema real.



Integrantes: Filipe Caldas de Souza de Oliveira

Professor Orientador: Anna Helena Reali Costa (PCS)

Co-orientador: Bruno Augusto Angélico (PTC)