

Comparação de técnicas de aprendizagem por reforço aplicadas no Super Mario Bros

Andrey Morais	Davi Aragão	Tiago Aroeira
<i>Instituto Tecnológico da Aeronáutica (ITA)</i>	<i>Instituto Tecnológico da Aeronáutica (ITA)</i>	<i>Instituto Tecnológico da Aeronáutica (ITA)</i>

1. Contextualização

Desde as primeiras inteligências artificiais (IAs), os Toy Problems, problemas sem interesse científico como jogos, são utilizados como forma de ilustrar os potenciais dos agentes inteligentes. Seu uso é derivado do fato desses problemas oferecerem um ambiente simplificado que facilita a visualização do comportamento do agente.

A primeira aplicação em jogos ocorreu em 1956, no mesmo ano do workshop de Dartmouth que estabeleceu a IA como um campo de pesquisa, pelo pesquisador Arthur Samuel com seu Game of Checkers.

Em 2016, o agente inteligente AlphaGo, desenvolvido pela DeepMind Technologies, venceu uma partida de Go contra o melhor jogador do mundo, quebrando o paradigma da superioridade da humana em jogos. A partir daí, outros agentes apareceram, como OpenIA Five, desenvolvido pela OpenIA para DoTA 2, que derrotou o time campeão do mundo nesse jogo em 2019.

Os jogos citados são conhecidos por sua complexidade, com possibilidades gigantescas de estratégias e combinações de ações. O avanço da eficácia nesse tipo de ambiente se deu com o advento de técnicas de treinamento por reforço, no qual o agente inteligente aprende sozinho a interagir no ambiente. Essa aprendizagem provém da busca do agente por maximizar a função de avaliação, função essa que compensa ou penaliza o comportamento adotado.

Dentro desse contexto, o trabalho atual propõe a implementação de duas das técnicas mais atuais aprendizagem por reforço, PPO e SAC, e a comparação do seu desempenho no jogo Super Mario Bros executado em um simulador do console Nintendo 8 Bits.

2. Objetivos

Implementar, treinar as redes PPO e SAC e comparar o seu desempenho na primeira fase do jogo Super Mario Bros executado em um simulador de console Nintendo 8 Bits.

3. Metodologia

1. Implementar uma API de comunicação com emuladores de jogos eletrônicos para obter dados do

comportamento e operação do jogo eletrônico Super Mario Bros.

2. Implementar e treinar o agente inteligente utilizando as redes PPO (Preferred Provider Organization) e SAC (Soft Actor-Critic).
3. Coletar e comparar os resultados de ambas as redes

3.1. Forma de análise

As duas redes serão comparadas quanto a seu desempenho: tempo de treinamento; e performance obtida no jogo.