Deep Reinforcement Learning

Fabrício Barth

May 11, 2022

1 Introdução

Segundo [1], desenvolver agentes que aprendem a atuar em um ambiente de alta dimensionalidade sempre foi um desafio para soluções baseadas em aprendizagem por reforço(RL). Até 2013, a maioria das aplicações de RL operavam nestes domínios com base em atributos determinados manualmente pelo projetista.

Em [1] os autores do artigo propõe uma variante do algoritmo Q-Learning [2] onde os pesos de uma rede neural são treinados no lugar de uma Q-table.

2 Algoritmo Q-Learning

Na figura 1 é apresentado o pseudo-código do algoritmo Q-Learning. Neste pseudo-código é possível ver como os pares $Q(s, a)$ são atualizados repetidas vezes através nas inúmeras interações do agente com o ambiente.

3 Algoritmo Deep Q-learning com Experience Replay

Este algoritmo foi proposto em [1] e tem o seguinte pseudo-código.

References

[1] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin A. Riedmiller. Playing atari with deep reinforcement learning. *CoRR*, abs/1312.5602, 2013.

Algorithm 1 Algoritmo Q-Learning

function *Q-Learning(env, α, γ, ϵ, ϵ_{min}, ϵ_{dec}, episódios)*
inicializar os valores de $Q(s, a)$ arbitrariamente
for todos os episódios **do**
    inicializar s a partir de env
    repeat
        $a \leftarrow escolha(s, \epsilon)$
        $s', r \leftarrow$ executar a ação a no env
        $Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$
        $s \leftarrow s'$
    until s ser um estado final
    if $\epsilon > \epsilon_{min}$ **then** $\epsilon \leftarrow \epsilon \times \epsilon_{dec}$
end for
return Q

Algorithm 2 Algoritmo Deep Q-Learning

Inicializa a memória D com capacidade N

[2] Christopher J. C. H. Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3):279–292, May 1992.