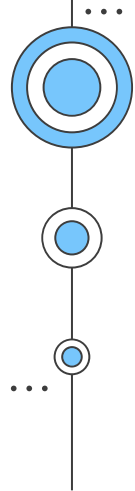
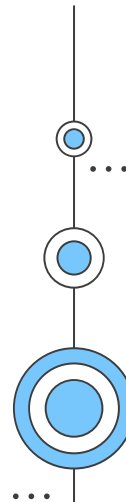


Aprendizado por Reforço



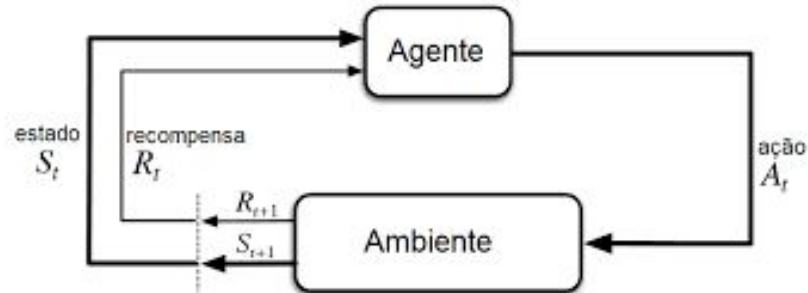
01

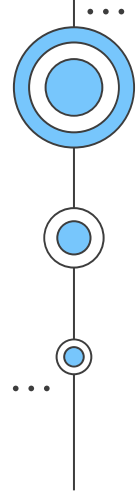
Introdução



Introdução

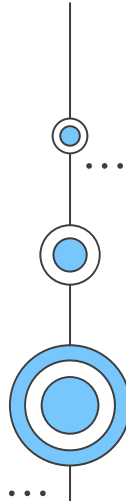
- Também conhecido como RL (Reinforcement Learning);
- Aprendizado por tentativa e erro (dados obtidos pela experiência);
- O modelo tenta aprender um comportamento para maximizar sua performance;
- Exemplo de métodos: Q-Learning, SARSA, R-Learning...





02

Conceitos



Conceitos



Agente

É a parte responsável por tomar decisões.



Ambiente

Mundo com o qual o Agente interage.



Estado

Descrição das condições em um instante.



Ação

Comando para interagir com o ambiente.



Recompensa

Feedback do ambiente pela ação tomada



Política

Próxima ação a ser tomada com base no estado atual.

Agente

- É o nosso software;
- Responsável por tomar decisões e aprender as melhores ações;
- No Xadrez, poderíamos pensar como se fosse o peão, mas na realidade é o jogador em si.



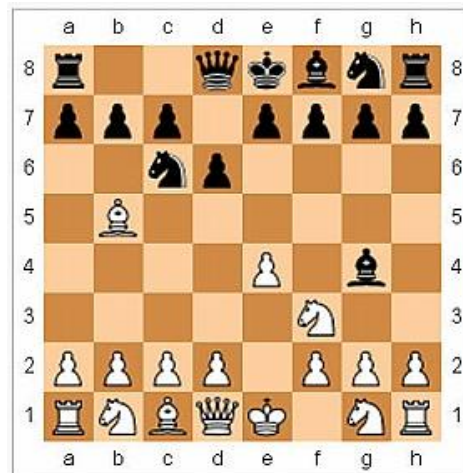
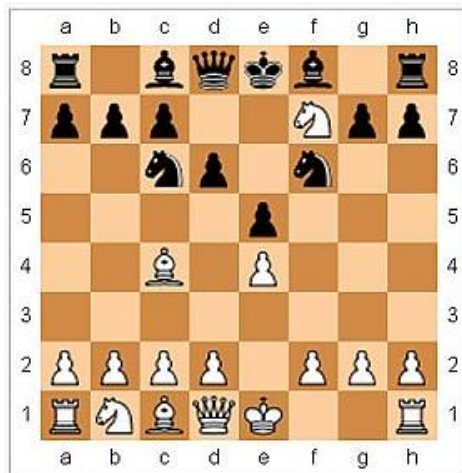
Ambiente

- Mundo com o qual o Agente interage.
- É o espaço que representa nosso problema, transmitindo informações ao Agente.
- Vai além do espaço "físico", também inclui personagens, sujeitos, seres...
- No Xadrez, além do tabuleiro as peças também fazem parte do ambiente.



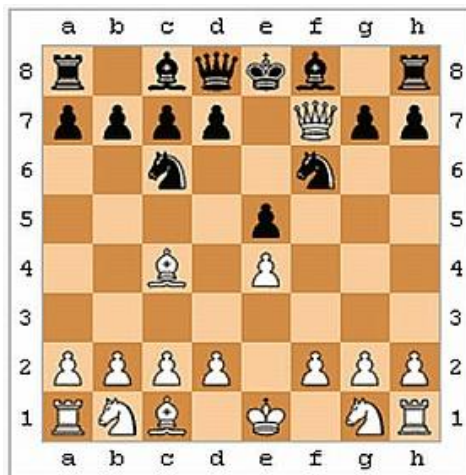
Estado

- É a descrição em um instante das condições do Agente e do Ambiente.
- É a partir dele que o Agente deve tomar suas decisões.
- No Xadrez, a diferente distribuição das peças configuram estados diferentes do jogo.



Ação

- Comando que o Agente escolhe para interagir com o ambiente.
- O espaço de ação é o conjunto de todos os movimentos possíveis.
- No Xadrez, a ação seria equivalente a um movimento e o espaço de ação a todos.



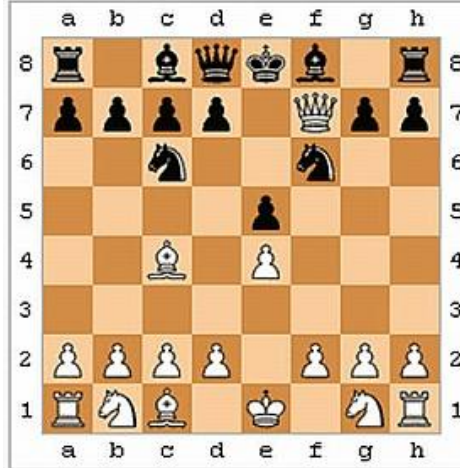
Recompensa

- A cada ação tomada, o Ambiente devolve um feedback ao Agente.
- Pode ser **positiva**, **negativa**, ou **nula**.
- No Xadrez, poderíamos capturar outra peça, ser capturado ou não acontecer nenhum dos casos.



Política

- A política é o que guia as escolhas do nosso agente, fornece qual a próxima ação a ser tomada com base no estado atual.
- O objetivo do nosso agente é encontrar a política ótima que escolhe a melhor ação para cada estado.
- No Xadrez, poderíamos considerar como política ótima a ação de capturar uma peça e não ser capturado por outra.



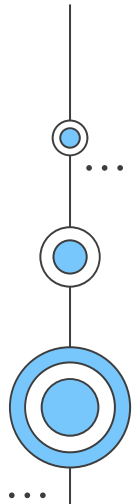
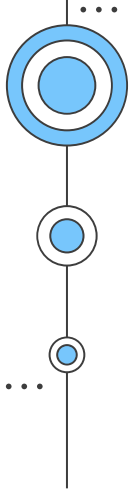
Resumo

- O Agente
- O Ambiente
- O estado
- A ação
- A recompensa



03

Referências



Referências

- <https://github.com/turing-usp/Workshop-de-Aprendizado-por-Reforco>
- <https://www.deeplearningbook.com.br/o-que-e-aprendizagem-por-reforco/>
- <https://www.imdb.com/title/tt6700846/>