



# Aplicação de inteligência artificial em jogos de estratégia complexos

Guilherme Vilar Balduino 743546  
Roberto Yamamoto 743594

# Objetivo



Explicar de maneira geral o que é necessário para uma inteligência artificial (IA) ser capaz de atuar em situações complexas como jogos de estratégia.

Com foco no jogo Starcraft, um jogo de estratégia em tempo real(RTS).

# Contexto

---

Em janeiro de 2019, a inteligência artificial da empresa DeepMind, AlphaStar, derrotou jogadores profissionais de Starcraft em uma série de onze partidas. Uma IA já superou anteriormente humanos profissionais em outros jogos de estratégia como xadrez, que é feito em sistema de turnos. A grande dificuldade de criar uma IA para um jogo de estratégia em tempo real e o grande número de fatores que influenciam na dinâmica do jogo.



# Como funciona Starcraft?

A dinâmica do jogo se altera de acordo com as diferentes estratégias para cada raça presente no jogo(Terran, Zerg e Protoss), conhecimento parcial do cenário, tal que o jogador somente tem conhecimento das regiões próximas aos seus edifícios e unidades, e os diversos tipos de unidades, tais quais tem diferentes utilidades, como os coletores(VCEs, Zangões e Sondas) e unidades aéreas(Viking, Corruptor, Fênix).



# Como funciona Starcraft?



# Aprendizado por reforço



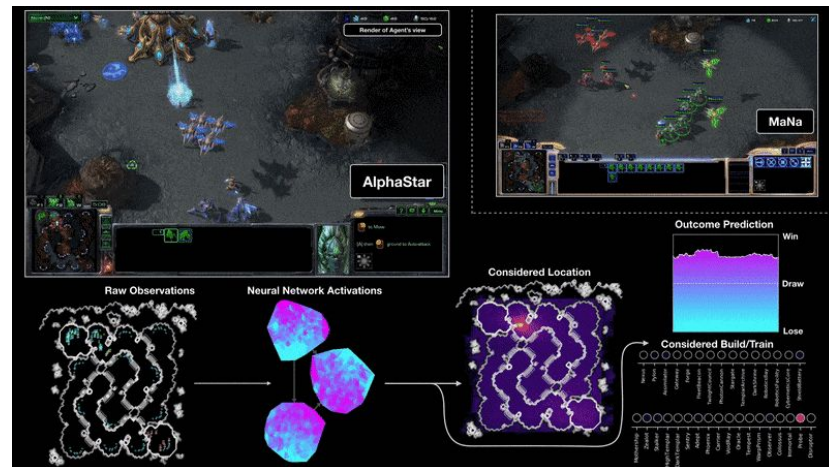
Aprendizado inspirado em como os humanos e animais aprendem a tomar decisões, com base a ganhos positivos e negativos.

Usando como base um momento em que o agente se encontra em um determinado estado, ele decide realizar uma ação específica dentre todas disponíveis no momento e recebendo imediatamente uma recompensa em resposta a isso.

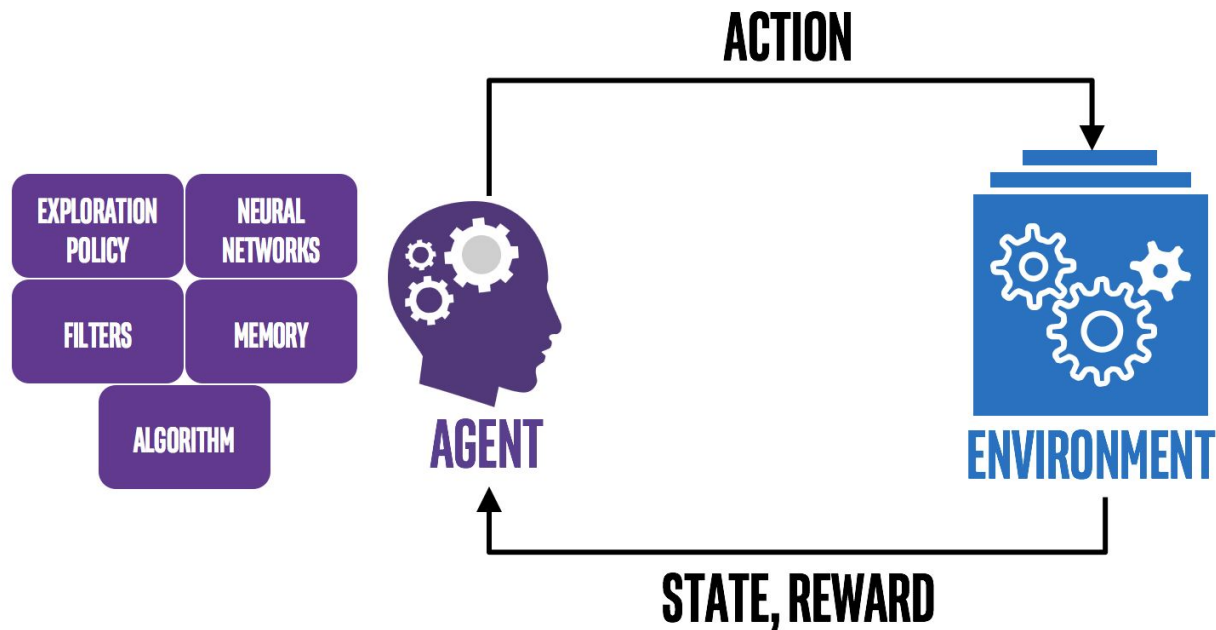
# Aprendizado por reforço

Isso pode ser visto em Starcraft como por exemplo: o agente pode criar unidades de combate, criar novos edifícios, expandir sua base, dentre outras várias possibilidades.

Em média, um jogador profissional realiza cerca de 400 movimentos por minuto.



# Aprendizado por reforço





# Principais elementos



- Política

Como o agente deve se comportar em determinada situação?

Representa as regras do ambiente

- Função Recompensa

Qual o objetivo a ser atingido atualmente?

Geralmente aplicada para tomar decisões de curto prazo?

# Principais elementos



- Função Valor

Qual a recompensa que o agente pode acumular a partir de um estado?

Assim sendo capaz de prever em qual estado a recompensa será maior. É utilizada para tomada de decisões de longo prazo

- Modelo do Ambiente

Uma representação de ambiente no qual o agente foi inserido e lista de ações que o agente é capaz de realizar

# Definições matemáticas



Valor das ações para escolher qual deve ser realizada em um momento específico com o intuito de calcular o quão boa é a ação.

$$Q_t(a) = \frac{R_1 + R_2 + \dots + R_{K_a}}{K_a}$$

Soma das recompensas obtidas com a ação e  $K_a$  representa a quantidade de vezes que a ação foi escolhida.

Com essa função definida, são determinados métodos para escolher as ações que serão tomadas em algum momento do problema. As abordagens mais comuns são a Greedy, e-greedy e softmax .

# Definições matemáticas



O agente tem o objetivo de maximizar a recompensa a longo prazo, de modo que é possível definir uma função que descreva o retorno, onde  $T$  é a última ação a ser tomada para solucionar o problema e  $G_t$  é a soma das recompensas recebidas depois do instante  $t$ , tudo isso somado ao conceito de desconto, sendo este uma taxa com valor entre 0 e 1, que determina o valor atual de futuras recompensas, assim fazendo o agente tomar decisões a curto ou a longo prazo. Assim criando a função:

$$G_t = R_{t+1} + \gamma^2 R_{t+2} + \gamma^3 R_{t+3} + \dots + \gamma^n R_T = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

# Referências:



<https://upcommons.upc.edu/handle/2117/120692>

<https://monografias.ufrn.br/jspui/handle/123456789/4322>

<https://nervanasystems.github.io/coach/>

<https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/>

<https://starcraft2.com/pt-br/>