

Inteligência Artificial - Notas de aula

Raoni F. S. Teixeira

Aula 1 - Agente Inteligente

1 Introdução

Este curso trata de sistemas que percebem e atuam de maneira racional em um ambiente. Esse tipo de sistema age de forma otimizada, ou seja, ele escolhe a melhor ação com base nos sinais que recebe. Vamos entender isso melhor com o conceito de agente inteligente, que veremos a seguir.

2 Agente Inteligente

Russell e Norvig [RN09] colocam o conceito de agente racional no centro da Inteligência Artificial. Um *agente*, como mostrado na Figura 1, é um sistema que *percebe* o ambiente com sensores e *age* sobre ele usando atuadores. Sensores podem ser câmeras ou detectores, enquanto atuadores são dispositivos que modificam o ambiente.

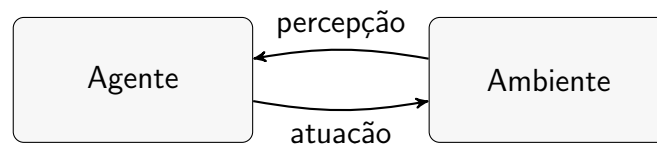


Figura 1: Agente interagindo com um ambiente.

Podemos imaginar um agente como um sistema que recebe informações do ambiente (percepções) e decide o que fazer (ações). Formalmente, isso é representado por uma função matemática chamada f , que transforma as percepções em ações:

$$f : \mathcal{P} \mapsto \mathcal{A}. \quad (1)$$

O desempenho do agente depende da interpretação dos sinais e das ações escolhidas. Um agente é considerado inteligente se suas ações maximizam o sucesso esperado. Avaliamos isso com uma função de desempenho g , que atribui uma nota às combinações de percepções e ações:

$$g : \mathcal{P} \times \mathcal{A} \mapsto [-\infty, \infty]. \quad (2)$$

O agente inteligente é uma função f que maximiza a expectativa de g :

$$\arg\text{-max}_f E[g(\mathcal{P}, f(\mathcal{P}))]. \quad (3)$$

Por exemplo, pense em um robô aspirador. Ele usa sensores para detectar sujeira e obstáculos e toma decisões sobre quando aspirar ou desviar. Se ele limpa bem a casa, digamos que suas decisões foram eficazes.

3 Panorama e Organização do Curso

Agora que entendemos o conceito de agente inteligente, você pode se perguntar: como essa ideia evoluiu até aqui? Vamos explorar a história fascinante desse campo.

A Figura 2 apresenta os principais marcos da IA, desde a criação do termo por John McCarthy em 1956 até o lançamento do ChatGPT em 2022.

O termo *Inteligência Artificial* foi introduzido por John McCarthy na conferência de Dartmouth, em 1956, marcando o início formal do campo. Pouco depois, em 1958, surgiu o Perceptron, o primeiro modelo de rede neural, e, em 1959, o algoritmo Minimax, que estudaremos na aula 4.

A retropropagação do erro (*backpropagation*), popularizada em 1986 por Geoffrey Hinton e colaboradores, revolucionou o treinamento de redes neurais, tornando possível o desenvolvimento de redes profundas, como veremos na aula 9.

Em 1997, a vitória do supercomputador Deep Blue da IBM sobre o campeão mundial de xadrez Garry Kasparov destacou a capacidade da IA em resolver problemas complexos. Em 2012, a rede neural AlexNet, desenvolvida por Alex Krizhevsky e colegas, conquistou a competição ImageNet, transformando o campo da visão computacional e estabelecendo os fundamentos das redes convolucionais modernas, que serão abordadas na aula 12.

Em 2016, o AlphaGo, da DeepMind, superou o campeão mundial de Go, Lee Sedol, em um jogo de alta complexidade devido ao vasto espaço de possibilidades, demonstrando avanços significativos no aprendizado de reforço – tema da aula 12. Finalmente, em 2022, o lançamento do ChatGPT pela OpenAI representou um marco nos modelos de linguagem natural, com

uma capacidade de gerar texto coerente e interagir de forma humana. A técnica subjacente, denominada *Transformer*, será estudada na aula 13.

Entre 1997 e 2012, avanços em hardware e software de processamento, como GPUs e arquiteturas CUDA (lançadas em 2007), juntamente com a disponibilização de fotos e textos em blogs e redes sociais, e a catalogação e publicação de bases de dados, criaram um ambiente para a aplicação de redes neurais e o desenvolvimento de novas abordagens.

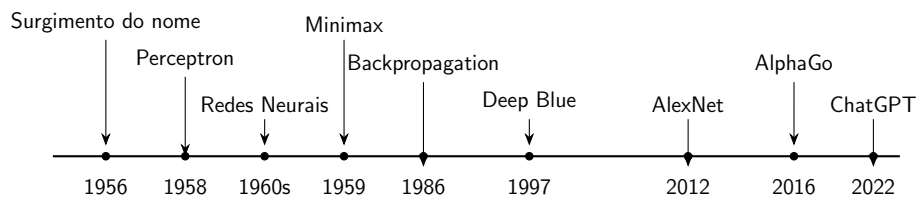


Figura 2: Linha do tempo com os principais marcos da inteligência artificial.

4 Discussão

Resolver problemas reais exige mais do que cálculos: precisamos fazer perguntas importantes. As percepções são precisas? O agente está interpretando os sinais corretamente? Há viés em suas decisões? Essas questões são técnicas, mas também éticas e sociais [CC16].

Além disso, problemas complexos como a redução da criminalidade, conhecidos como *wicked problems*, não têm soluções definitivas. Cada resposta reflete a visão de mundo de quem a propõe [RW73].

Referências

- [CC16] Kate Crawford and Ryan Calo. There is a blind spot in ai research. *Nature*, 538:311–313, October 2016.
- [RN09] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall Press, USA, 3rd edition, 2009.
- [RW73] Horst W. J. Rittel and Melvin M. Webber. Dilemmas in a general theory of planning. *Policy Sciences*, 4(2):155–169, June 1973.