

# Introdução à IA – Conceito, Histórico e Fundamentos

Profa. Mercedes Gonzales  
Márquez

---

# Tópicos

- Conceito de IA
- Histórico de IA
- Fundamentos de IA

# Conceito

- Historicamente, os pesquisadores tem seguido diversas versões diferentes de IA. Alguns tem definido inteligência como “fidelidade ao desempenho humano”, enquanto outros preferem uma definição abstrata e formal chamada de “racionalidade”.

# Conceito

- Os métodos usados são diferentes:
- Por um lado a busca da inteligência semelhante à humana deve ser em parte uma ciência empírica relacionada à psicologia, envolvendo observações e hipóteses sobre o comportamento humano real.
- Por outro lado, a abordagem racionalista envolve uma combinação de matemática e engenharia que se conecta à estatística, teoria de controle e economia.

# Conceito

- Dessas duas dimensões: humano contra racional, comportamento contra pensamento, existem quatro combinações possíveis.

**Sistemas que agem como humanos**

Sistemas que pensam como humanos

Sistemas que pensam racionalmente

Sistemas que agem racionalmente

# Conceito

- I. Sistemas que agem como seres humanos:  
“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.”  
(KURZWEIL, 1990).
  - Essa definição foi proposta por Alan Turing, um dos pioneiros da IA quem argumentou que uma máquina pode ser considerada inteligente se ela for capaz de agir de forma indistinguível de um ser humano.
  - A definição de Turing é baseada na ideia de que a inteligência humana pode ser medida pela capacidade de realizar tarefas que normalmente exigiriam inteligência humana, como:

# Conceito

- Entender linguagem
- Reconhecer imagens e objetos
- Resolver problemas complexos
- Aprender e se adaptar a novas situações

Turing propôs o "Teste de Turing" como uma forma de avaliar se uma máquina é capaz de agir de forma humana. O teste consiste em um julgador que conversa com uma máquina e um ser humano, sem saber quem é quem. Se o julgador não for capaz de distinguir a máquina do ser humano, então a máquina é considerada capaz de agir de forma humana.

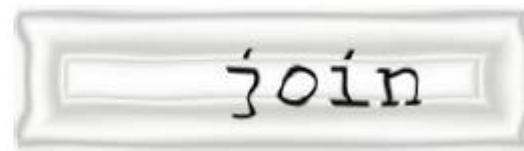
# Conceito

Exemplo: CAPTCHA = *Completely Automated Public Turing test to Tell Computers and Humans Apart.*

Um teste de reconhecimento de palavras em imagens, supostamente simples para humanos.

Feito para identificar programas que tentam utilizar serviços na Internet, como abrir contas de e-mail grátis.

É um teste de Turing, tenta-se identificar uma máquina passando por um ser humano.



# Conceito

- II. Sistemas que pensam como seres humanos:  
“O novo e interessante esforço para fazer os computadores pensarem... máquinas com mentes, no sentido total e literal”. (HAUGELAND, 1985).
  - Esta é uma definição mais ampla e profunda do que a definição de "agir de forma humana" proposta por Alan Turing. Essa definição foi proposta por John McCarthy quem argumentou que a IA não se limita apenas a realizar tarefas que normalmente exigiriam inteligência humana, mas também envolve a capacidade de pensar de forma humana. Trata-se então de uma estratégia de modelagem cognitiva.

# Conceito

- No entanto, essa definição também é mais difícil de medir e avaliar, pois envolve a capacidade de pensar de forma humana, o que é um processo complexo e multifacetado. Podemos aprender sobre o pensamento humano de três maneiras: Introspecção, experimentos psicológicos, imagens do cérebro.

# Conceito

- III. Sistemas que pensam racionalmente:

“O estudo das faculdades mentais pelo uso de modelos computacionais.” (CHARNIAK; MCDERMOTT, 1985).

- Um sistema de IA que "pensa racionalmente" é capaz de:  
Representar conhecimento e crenças de forma lógica e consistente, raciocinar e deduzir conclusões a partir de premissas e evidências.

# Conceito

- Essa definição é mais ampla e inclusiva do que as definições anteriores, pois não se limita apenas a realizar tarefas que normalmente exigiriam inteligência humana, mas também enfatiza a capacidade de pensar e raciocinar de forma lógica e racional.
- Além disso, essa definição também é mais fácil de medir e avaliar, por exemplo por meio de testes e experimentos.

# Conceito

- IV. Sistemas que atuam racionalmente:

“A Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (POOLE et al., 1998).

“O agente racional é aquele que age para obter o melhor resultado ou, se existe incerteza, o melhor resultado esperado” (Russel & Norvig, 2010)

- Esta é uma definição que enfatiza a capacidade de um sistema de IA de realizar ações e tomar decisões de forma racional e eficiente, e é mais focada na ação e na tomada de decisões do que na capacidade de pensar ou raciocinar.

# Fundamentos da IA

- Filosofia:
  - Regras formais podem ser usadas para obter conclusões válidas?
  - Como a mente (o intelecto) se desenvolve a partir de um cérebro físico?
  - De onde vem o conhecimento?
  - Como o conhecimento conduz a ação?

# Fundamentos da IA

- Filosofia:

- Regras formais podem ser usadas para obter conclusões válidas?

Aristoteles (384-322 aC) foi o primeiro a formular um conjunto preciso de leis que governam a parte racional da mente.

Regras formais pode ser usadas para obter conclusões válidas, porém podem ter limitações como complexidade e dificuldade para serem gerenciadas à medida que um sistema de IA cresce.

# Fundamentos da IA

Elas também podem não ser flexíveis o suficiente para lidar com situações imprevisíveis e ambíguas.

Essas regras formais não permitem que a IA aprenda com a experiência ou se adapte a novas situações.

# Fundamentos da IA

- Como a mente (o intelecto) se desenvolve a partir de um cérebro físico?

Segundo o dualismo proposto por Descartes há uma primeira distinção entre matéria e mente, pois uma concepção puramente física da mente deixa pouco espaço para o livre arbítrio. Há uma parte da mente humana (ou alma, ou espírito) que transcende a natureza, isenta de leis físicas.

Segundo o materialismo a operação do cérebro de acordo com as leis da física constitui a mente.

# Fundamentos da IA

- De onde vem o conhecimento?

O empirismo defende que o conhecimento vem da experiência sensorial e da observação do mundo ao nosso redor.

O positivismo lógico é o empirismo lógico que sustenta que o conhecimento pode ser caracterizado por teorias lógicas conectadas a sentenças de observação que correspondem a entradas sensoriais.

A teoria da confirmação sugere que as pessoas tendem a buscar e interpretar informações de maneira que confirme suas crenças e expectativas prévias.

# Fundamentos da IA

- Como o conhecimento conduz a ação?

Utilitarismo: Teoria que defende que as ações devem ser avaliadas com base em seu impacto na felicidade ou no bem-estar geral. Uma ação é considerada correta se ela maximizar a felicidade ou o bem-estar geral.

Ética deontológica: defende que as ações devem ser avaliadas com base em sua conformidade com regras e deveres morais.

# Fundamentos da IA

- Matemática:
  - Quais são as regras formais para obter conclusões válidas?
  - O que pode ser computado?
  - Como raciocinamos com informações incertas?

# Fundamentos da IA

- Matemática:

- Quais são as regras formais para obter conclusões válidas?

A lógica formal começou com Boole que definiu a lógica proposicional ou lógica booleana que trabalha com valores booleanos e operações lógicas. Mais adiante surgiu a lógica de primeira ordem que trabalha com objetos, relações e quantificadores.

# Fundamentos da IA

- Como raciocinamos com informações incertas?

A teoria de probabilidade pode ser vista como uma lógica generalizadora para situações com informações incertas.

Bayes propôs uma regra para atualizar probabilidades à luz de novas evidências.

A formalização da probabilidade, combinada com a disponibilidade de dados levou ao surgimento da estatística.

# Fundamentos da IA

- O que pode ser computado?

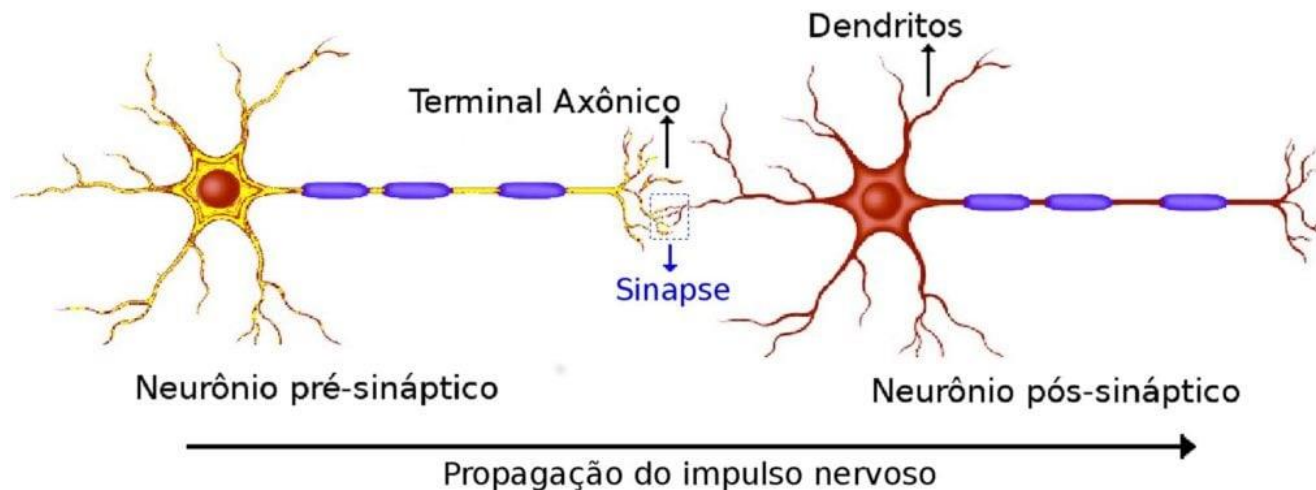
Conceito de tratabilidade: um problema é chamado de intratável se o tempo necessário para resolver instâncias dele cresce exponencialmente com o tamanho das instâncias. A teoria de NP-completude fornece uma base para analisar a tratabilidade dos problemas.

# Fundamentos da IA

- Neurociência:

Como o cérebro processa informações?

As funções cognitivas resultam da operação eletroquímica dos neurônios. Ou seja, uma coleção de células simples pode levar ao pensamento, à ação e à consciência.



# Fundamentos da IA

Um neurônio faz conexões com 10 a 100 mil outros neurônios em junções chamadas “sinapses”. Os sinais se propagam por meio de uma complicada reação eletroquímica.

# Fundamentos da IA

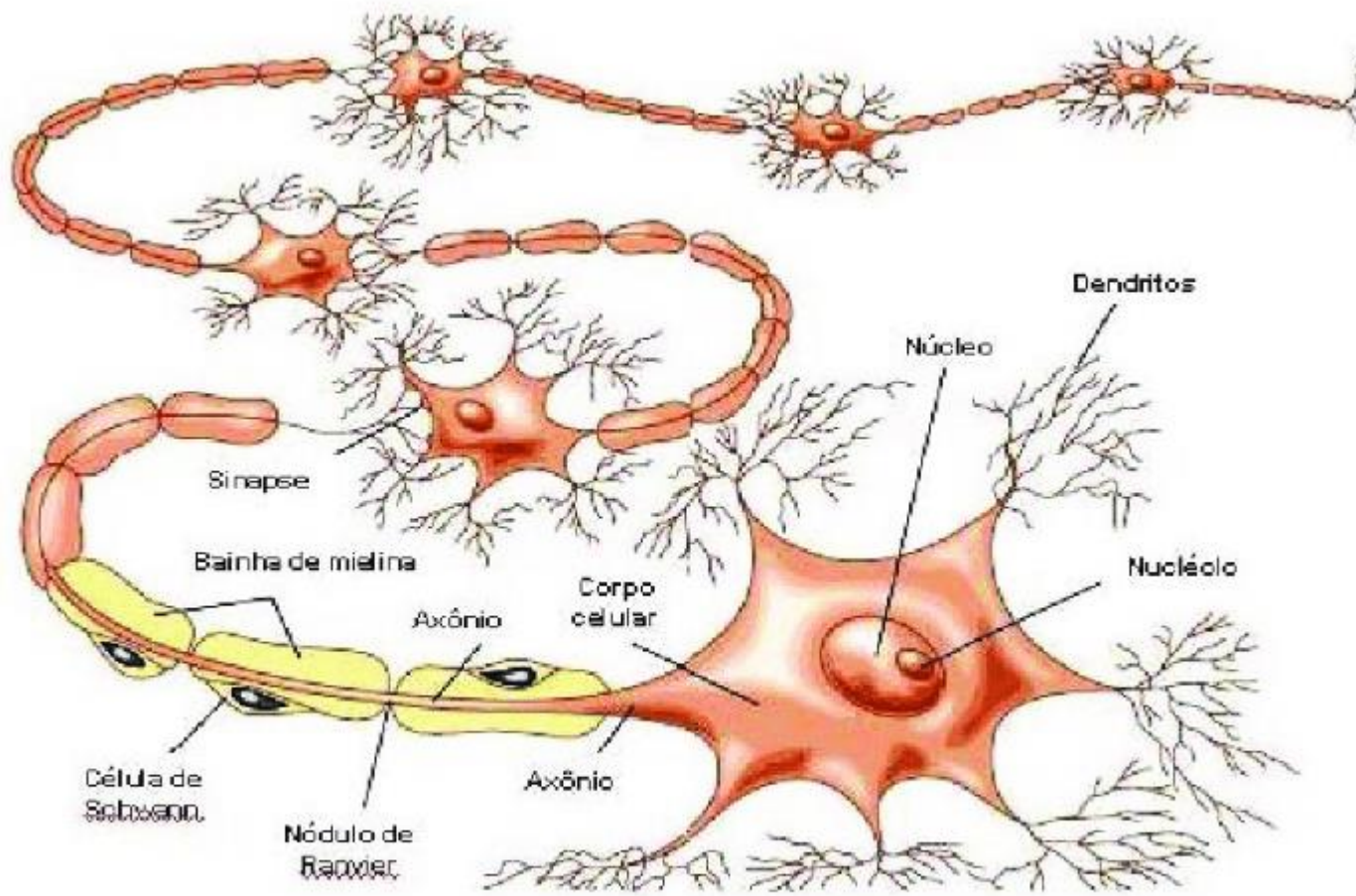


Ilustração de neurônios conectados, mostrando a transmissão de informações em uma rede neuronal.

# Fundamentos da IA

O desenvolvimento de interfaces cérebro-máquina para detecção e controle motor apresenta uma possibilidade de restauração da função de indivíduos com deficiência.

# Fundamentos da IA

- Psicologia e Linguística:

Como os seres humanos e os animais pensam e agem?

Como a linguagem se relaciona com o pensamento?

Ler páginas 12 e 15 do livro Inteligência Artificial Uma abordagem moderna. 4ta edição. Stuart Russell e Peter Norvig.

# Riscos e Benefícios da IA

- Benefícios:
  - Se tivermos acesso a uma inteligência mecânica substancialmente maior, o teto das nossas ambições será substancialmente elevado.
  - A humanidade pode ser liberta do trabalho braçal e repetitivo, aumentando a produção de bens e serviços.
  - Prováveis curas para doenças e soluções para mudanças climáticas e a escassez de recursos.

# Riscos e Benefícios da IA

- Riscos:
  - Armas autônomas letais
  - Vigilância e persuasão.
  - Tomada de decisão tendenciosa.
  - Impacto no emprego
  - Aplicações de segurança crítica
  - Cibersegurança

# Riscos e Benefícios da IA

- Riscos:
  - Os sistemas de IA passam a assumir uma parcela maior de funções sociais anteriormente desempenhadas por seres humanos. Assim como os seres humanos usaram essas funções no passado para causar prejuízo, pode-se esperar que os humanos usem mal os sistemas de IA para causar ainda mais prejuízo.
  - Uma superinteligência artificial poderia ser uma má ideia de acordo com Turing:

# Riscos e Benefícios da IA

“ Parece provável que, quando o método de raciocínio de máquina for iniciado, não levará muito tempo para que ultrapasse nossos fracos poderes (...) Portanto, em algum estágio poderemos esperar que as máquinas tomem o controle ...” (Turing).

Os valores ou objetivos colocados na máquina devem ser alinhados aos do ser humano. Consequências negativas surgirão de um sistema implantado com um objetivo incorreto. Além disso, quanto mais inteligente for o sistema, mais negativas serão as consequências.