

Inteligência Artificial

Visão Geral

Prof. Dr. Tiago Araújo

O que é IA?

O que é IA?

A automação de atividades que associamos ao pensamento humano, atividades como a tomada de decisões, resolução de problemas, aprendizagem... (Bellman, 1978)	O estudo das faculdades mentais através do uso de modelos computacionais (Charniak+McDermott, 1985)
O estudo de como fazer com que os computadores façam coisas em que, no momento, as pessoas são melhores" (Rich+Knight, 1991)	"O ramo da informática que se preocupa com a automação do comportamento inteligente" (Luger+Stubblefield, 1993)

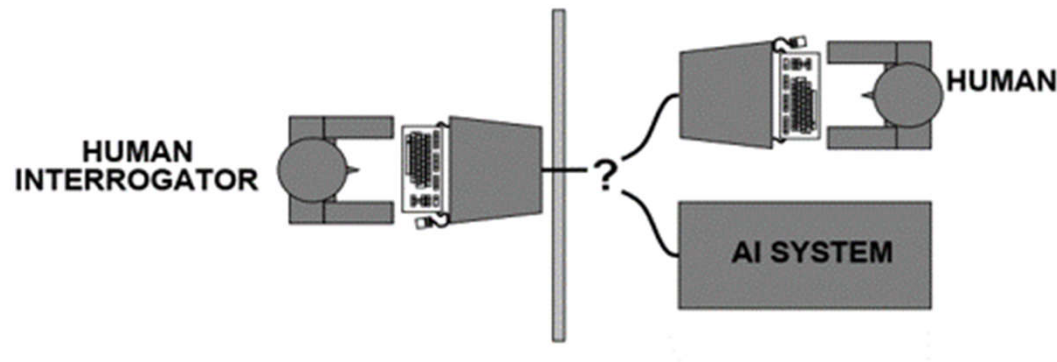
As visões de IA se encaixam em quatro categorias:

Pensando humanamente	Pensando racionalmente
Agindo humanamente	Agindo racionalmente

Examinando eles, vamos optar por agir racionalmente (mais ou menos)

Agindo Humanamente: O Teste de Turing

- Turing (1950) "Máquinas de Computação e Inteligência":
 - ❖ "As máquinas podem pensar? ➡ "As máquinas podem se comportar de forma inteligente?"
 - ❖ Teste operacional para comportamento inteligente: o jogo de imitação



Agindo Humanamente: O Teste de Turing

- ❖ Previa que em 2000, uma máquina poderia ter 30% de chance de enganar uma pessoa leiga por 5 minutos
- ❖ Antecipou todos os principais argumentos contra a IA nos 50 anos seguintes

Problema: O teste de Turing não é reproduzível, construtivo ou receptivo à análise matemática

Pensando Humanamente: Ciência Cognitiva

- 1960 “Revolução cognitiva”: A psicologia do processamento da informação foi substituída prevalecendo a ortodoxia do comportamentalismo dominante
- Requer teorias científicas sobre as atividades internas do cérebro
 - ❖ Qual o nível de abstração? "Conhecimento" ou "circuitos"?
 - ❖ Como validar? Requer:
 - 1) Previsão e comportamento de teste de sujeitos humanos (de cima para baixo) ou
 - 2) Identificação direta a partir de dados neurológicos (baixo para cima)
- Ambas as abordagens (aproximadamente, Ciência Cognitiva e Neurociência Cognitiva) são agora distintos da IA

Pensando Racionalmente: Leis do Pensamento

- **Normativa** (ou **prescritiva**) em vez de **descritiva**
- Aristóteles: quais são os argumentos/processos de pensamento corretos?
- Várias escolas gregas desenvolveram várias formas de lógica: **notação e regras de derivação** do pensamento; podem ou não ter procedido à ideia de mecanização.
- Linha direta através da matemática e da filosofia para a IA moderna.
- Problemas:
 - 1) Nem todo comportamento inteligente é mediado por uma deliberação lógica.
 - 2) Qual é o propósito de pensar? Que pensamentos eu deveria ter?

Agindo Racionalmente

- Comportamento **Racional**: Fazer a coisa certa
- A coisa certa: é o que se espera que maximize o alcance das metas, dadas as informações disponíveis.
- Não envolve necessariamente o pensamento - por exemplo, reflexo de piscar - mas o pensamento deve estar a serviço da ação racional
- Aristóteles (Ética Nicomacheana)

Toda arte e toda investigação, e igualmente toda ação e perseguição, é pensada para visar algum bem.

Fundações de IA

- **Filosofia**
 - Como surge a mente de um cérebro?
 - Conhecimento?
- **Matemática**
 - Quais as regras formais para gerar conclusões?
 - Incerteza?
- **Economia**
 - Como tomar decisões baseado em preferências?
- **Neurociência**
 - Como nós processamos informação?
- **Psicologia**
 - Como humanos e animais agem?
- **Linguística**
 - Como idioma se relaciona com o pensamento?
- **Computadores**
 - Como criar computadores e dispositivos para inteligência?

	Supercomputer	Personal Computer	Human Brain
Computational units	10^6 GPUs + CPUs 10^{15} transistors	8 CPU cores 10^{10} transistors	10^6 columns 10^{11} neurons
Storage units	10^{16} bytes RAM 10^{17} bytes disk	10^{10} bytes RAM 10^{12} bytes disk	10^{11} neurons 10^{14} synapses
Cycle time	10^{-9} sec	10^{-9} sec	10^{-3} sec
Operations/sec	10^{18}	10^{10}	10^{17}

Agentes Racionais

- Um **agente** é uma entidade que percebe e age
- Esta disciplina é sobre a concepção de agentes racionais
- Abstratamente, um agente é uma função que vai desde a percepção de histórias até as ações:

$$f : \mathcal{P}^* \rightarrow \mathcal{A}$$

- Para qualquer classe de ambientes e tarefas, procuramos o agente (ou classe de agentes) com o melhor desempenho.
- Caveat: limitações computacionais tornam a racionalidade perfeita inalcançável ➡ projeto do melhor **programa** para determinados recursos da máquina.

Pré-história da IA

- Filosofia: lógica, métodos de raciocínio mental como fundamentos do sistema físico de aprendizagem, linguagem, racionalidade.
- Matemática: representação formal e cálculo de algoritmos de prova, (in)decidibilidade, (in)tratabilidade, probabilidade
- Psicologia: adaptação, fenômenos de percepção e controle motor, técnicas experimentais (psicofísica, etc.)
- Linguística: representação do conhecimento, gramática
- Neurociência: substrato físico para a atividade mental
- Teoria de controle, sistemas homeostáticos, estabilidade, projetos simples de agentes ideais

História resumida da IA

- 1943 : McCulloch, Pitts: Modelo de circuito booleano do cérebro
- 1950 : Turing's ``Computing Machinery and Intelligence'' (Maquinaria e Inteligência da Turing)
- 1952-69 : Olhe, mãe, sem mãos!
 - 1950 : Programas iniciais de IA, incluindo o programa de damas de Samuel, Newell & Simon's Logic Theorist, Motor de Geometria da Gelernter
 - 1956 : Reunião de Dartmouth: "Inteligência Artificial" adotada
 - 1965 : Algoritmo completo de Robinson para raciocínio lógico

História resumida da IA

- 1966-74 : IA descobre complexidade computacional, a pesquisa da rede neural quase desaparece
- 1969-79 : Desenvolvimento precoce de sistemas baseados no conhecimento
- 1980-88 : A expansão setor de sistemas especializados
- 1988-93 : Bustos da indústria de sistemas especializados:
Inverno AI
 - 1985-95 : As redes neurais voltam à popularidade
 - 1988- : Ressurgimento de métodos probabilísticos e teóricos de decisão, Rápido aumento da profundidade técnica da IA convencional, ``Nouvelle AI": ALife, GAs, computação leve.

O Estado da Arte

- Qual dos seguintes aspectos pode ser feito no momento?
 - ❖ Jogar um jogo decente de tênis de mesa
 - ❖ Conduzir ao longo de uma estrada curva de montanha
 - ❖ Dirigir no centro do Cairo
 - ❖ Jogar um jogo de StarCraft II a nível competitivo
 - ❖ Descobrir e provar um novo teorema matemático
 - ❖ Escrever uma história intencionalmente engraçada
 - ❖ Prestar assessoria jurídica competente em uma área especializada do direito
 - ❖ Traduzir o inglês falado para o sueco falado em tempo real
 - ❖ Prever estrutura de proteínas
 - ❖ Escrever TCCs e artigos científicos

Riscos e benefícios

- + Menos tarefas repetitivas
- + Aumento da produção e serviços
- + Acelerar pesquisa
- + Cura de doenças
- + Impacto ambiental
- + Solução para falta de recursos
- Armas autônomas letais
- Vigilância e persuasão
- Tomada de decisão enviesada
- Impactos na força de trabalho
- Aplicações críticas
- Cibersegurança