

Inteligência Artificial

Visão Geral & Agentes Inteligentes



O que é IA?

A automação de atividades que associamos ao pensamento humano, atividades como a tomada de decisões, resolução de problemas, aprendizagem...	O estudo das faculdades mentais através do uso de modelos computacionais
O estudo de como fazer com que os computadores façam coisas em que, no momento, as pessoas são melhores"	"O ramo da informática que se preocupa com a automação do comportamento inteligente"

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)

Prof. Dr. Tiago Araújo

Agindo Humanamente: O Teste de Turing

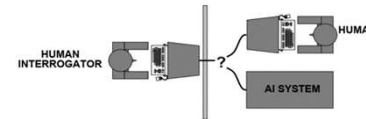


- Turing (1950) "Máquinas de Computação e Inteligência":
 - ❖ "As máquinas podem pensar?"
 - "As máquinas podem se comportar de forma inteligente?"
 - ❖ Teste operacional para comportamento inteligente: o jogo de imitação

❖ Prevê 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



2

Pensando Humanamente: Ciência Cognitiva

- 1960 "Revolução cognitiva": A psicologia do processamento da informação foi substituída prevalecendo a ortodoxia do comportamentalismo (Behaviorismo) 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?



3

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.

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.



4

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?

5

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

6

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

7

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 especialistas
- 1988-93: Avanços da indústria de sistemas especialistas: Inverno da IA
 - 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.
- 1997: Deep Blue da IBM vence campeão enxadrista Garry Kasparov
- 1998 – Yann Leccun publica o artigo *Gradient-based learning applied to document recognition*
- 2002 – iRobot lança a Roomba, um aspirador de pó automático
- 2009 – Primeiro carro autônomo do Google
- 2011:14 – Empresas lançam assistentes virtuais
- 2013 – A primeira CNN vencedora do ILSVRC
- 2014 – Primeira GAN criada
- 2016 – AlphaGo vence contra o campeão mundial de Go por 4–1.
- 2020 – OpenAI lança o GPT-3
- 2022 – GATO, um agente geral, é criado pela DeepMind

8

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



9

Riscos e benefícios

- | | |
|--|---|
| <ul style="list-style-type: none"> + Menos tarefas repetitivas + Aumento da produção e serviços + Acelerar pesquisa + Cura de doenças + Impacto ambiental + Solução para falta de recursos | <ul style="list-style-type: none"> – Armas autônomas letais – Vigilância e persuasão – Tomada de decisão enviesada – Impactos na força de trabalho – Aplicações críticas – Cibersegurança |
|--|---|

10

PAGE

- | | |
|---|---|
| <ul style="list-style-type: none"> • Deve primeiro especificar a configuração para o projeto de agente inteligente • Considere, por exemplo, a tarefa de projetar um táxi automatizado: • Perceptions - Percepções ?? • Actions - Ações ?? • Goals - Objetivos ?? • Environment – Ambiente ?? | <ul style="list-style-type: none"> • Deve primeiro especificar a configuração para o projeto de agente inteligente • Considere, por exemplo, a tarefa de projetar um táxi automatizado: • Percepções ?? Vídeo, acelerômetro, sensores do motor, teclado, GPS... • Ações ?? Volante, aceleração, freio, buzina, autofalante/tela • Objetivos ?? Segurança, alcançar o destino, maximizar lucros, cumprir leis, conforto do passageiro,... • Ambiente ?? Ruas urbanas dos EUA, rodovias, tráfego, pedestres, clima, clientes... |
|---|---|

Agente de compras da Internet?

11

Agente Racionais

- Sem perda de generalidade, "metas" especificáveis por **medida de desempenho** definindo um avaliador numérico para qualquer histórico ambiental
- **Ação racional**: qualquer ação que maximize o valor esperado da medida de desempenho **dada a sequência percebida até o momento**
- **Racional ≠ onisciente**
- **Racional ≠ clarividente**
- **Racional ≠ bem sucedido**

Funções e programas dos agentes

- ❖ Um agente é completamente especificado pelo mapeamento da função do agente percebe sequências de ações
 - ❖ Em princípio, é possível fornecer cada sequência possível para ver o que ela faz. Obviamente, uma tabela de pesquisa seria normalmente imensa.
- ❖ Uma função de agente (ou uma pequena classe de equivalência) é racional.
- ❖ Objetivo: encontrar uma maneira de implementar a função de agente racional de forma concisa

12

Tipos de Ambiente

	Solitário	Gamão	Internet Shopping	Taxi
Acessível??	Sim	Sim	Não	Não
Determinístico??	Sim	Não	Em parte	Não
Episódico??	Não	Não	Não	Não
Estático??	Sim	Meio	Meio	Não
Discreto??	Sim	Sim	Sim	Não

O tipo de meio ambiente determina em grande parte o projeto do agente.

O mundo real é (naturalmente) inacessível, estocástico, sequencial, dinâmico, contínuo

13

Funções e programas dos agentes

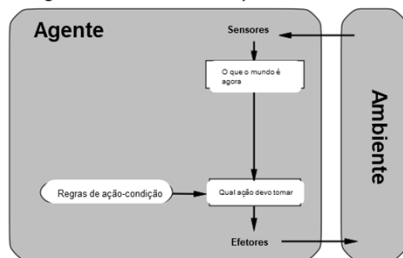
❖ Um programa de agente tem uma única percepção como entrada, mantém o estado interno:

```
function AGENTE-ESQUELETO(percepcao) returns acao
  static: memoria, a memoria do agente em relacao ao mundo

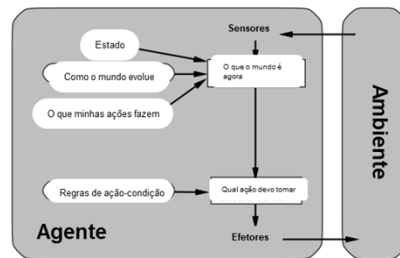
  memoria ← ATUALIZA-MEMORIA(memoria, percepcao)
  acao ← ESCOLHA-MELHOR-ACAO(memoria)
  memoria ← ATUALIZA-MEMORIA(memoria, acao)
  return acao
```

14

Agentes reflexivos simples

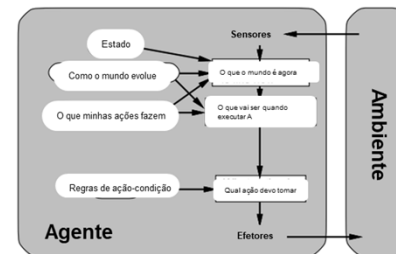


Agentes reflexivos com estados

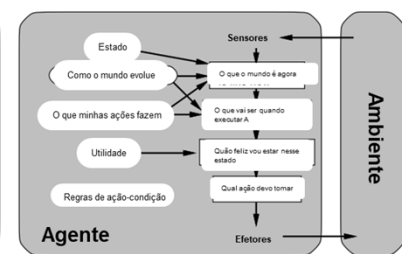


15

Agentes Baseados em Metas



Agentes Baseados em Utilidade



16