

4. IA: ESTUDO DA MENTE OU PROCESSAMENTO DE CONHECIMENTO?

- 1956: reunião que “cria” a IA:
 - estudo da mente: Newell
 - processamento de conhecimento: Feigenbaum
 - primeira preocupações: representação (McCarthy), resolução de problemas (Newell e Simon), modelagem da memória (Minsky, Miller), reconhecimento de padrões (Greenblatt)
- modelos de computação anteriores à IA: máquina de Turing (1936), sistema de produção (Post, 1936 e 1943), cálculo Lambda (Church, 1936 e 1941), funções recursivas gerais (Kleene, 1936), lógica combinatória (Curry, 1958), máquina de Mealey (1955), máquina de Kleene (1956), autômatos celulares (von Neumann, 1956), máquina de Moore (1956)
- modelo de computação de McCarthy (1958): linguagem Lisp
- modelos de computação associados à IA: sistemas formais (Smullyan, 1961), algoritmos genéticos (Holland, 1975), lógica de cláusulas Horn (Kowalski, 1974), nodos conexionistas (Rumelhart e McClelland, 1976), planejamento (Chapman, 1987), Improviso (Agre e Chapman, 1987)
- tese da IA forte:
 - a inteligência humana pode ser replicada
 - as explicações das funções cognitivas são apoiadas por processos computacionais
- tese da IA fraca:
 - a inteligência humana pode, no melhor dos casos, apenas ser simulada
 - a meta para se atingir a inteligência artificial consiste em projetar máquinas que sejam capazes de exibir um comportamento inteligente
- questões formuladas na seção 1 podem ser resumidas na seguinte (Hofstadter, 1979): um sistema cognitivo (sistema lógico-formal, teoria matemática, cérebro como “base” da mente) pode ocupar-se de si mesmo? Esta pergunta remonta aos problemas propostos por Hilbert (1900), Russel e Whitehead (*Principia Mathematica* tenta derivar toda matemática a partir da lógica) e outras questões relativas à provas matemáticas:
 - existe um método (definitivo), ou seja um processo mecânico, que possa ser aplicado à uma declaração matemática, o qual possa dizer se esta pode ser provada?
 - não:
 - Hardy (1928): afirma mas não prova formalmente
 - Turing (1936)
 - superioridade da mente frente à dispositivos mecânicos
 - determinismo: modelagem de problemas complexos a partir de elementos (máquina de estados, autômatos, etc.)
 - a matemática é completa e consistente?
 - não (Gödel, 1931) e o Teorema da Incompletude: se S for consistente (não contém um teorema que seja a negação de um outro), então existem sentenças verdadeiras que são teoremas em S e que não podem ser demonstradas:
 - é possível escrever declarações/sentenças matemáticas que se referem a elas mesmas (do tipo “Eu estou mentindo”, Esta sentença é falsa” ou “Esta declaração não pode ser provada”)
 - tais declarações não podem ser demonstradas ou seja provadas como sendo verdadeiras (levariam à uma contradição) ou como sendo falsas (idem)

- logo, uma declaração deste tipo não pode ser demonstrada a partir dos axiomas o que prova que, por exemplo, a aritmética (foi mostrado que esta pode ser formulada apenas em termos de manipulação de símbolos - números e operadores, ou seja usa-se números para fazer afirmações sobre números) não é completa no sentido proposto por Hilbert (cada declaração - por exemplo $2+2=4$ - pode ser provada como sendo V ou F)
- o exemplo usado por Gödel (“esta declaração não pode ser demonstrada”) é portanto uma declaração verdadeira (paradoxo ?) mas isto somente pode ser determinado por um observador externo (o próprio sistema axiomático não pode chegar à esta conclusão por si)
- o que isto tem a ver com a IA e a questão da mente?
 - um sistema formal S é um aparato simbólico (regras bem determinadas) o que implica em combinações simbólicas (teoremas de S) advindas das regras de S
 - a máquina de Turing é um sistema formal (Teorema de Gödel se aplica)
 - se o cérebro funcionar como uma máquina de Turing, permanecer consistente implica em uma limitação (por exemplo, não conseguirá demonstrar todas as sentenças verdadeiras da aritmética elementar)
 - mas, como o cérebro/mente (sistema S) consegue demonstrar (fora de S) que uma sentença (indemonstrável no sentido de Gödel) é de fato verdadeira, então o cérebro/mente não é uma máquina de Turing, e um corolário seria: cérebro/mente não “pensa” algoritmicamente (Lucas, 1961; Penrose, 1989)
- qual é sua opinião a respeito disto?
- Leitura:
 - Searle x Churchland and Churchland (*Sc.Am.* jan. 1990) [obrigatória]
 - The first AI debate (cap. 5 S. Franklin, *Artificial Minds*) [obrigatória]
 - Sloman (*Art. Intelligence* vol. 56: 355, 1992) [obrigatória pelo menos até seção 7]
 - Penrose (*The Emperor’s New Mind*, caps. 1, 2) [opcional]
 - Dreyfus and Dreyfus (cap. 2 S.R. Graubard, *The Artificial Intelligence Debate*) [opcional]
 - Lucas (Minds, Machines and Gödel. . *Philosophy* 36: 112, 1961) ou cap. XV de *Gödel, Escher and Bach* (Hofstadter) [opcional]
 - Whitely (Minds, Machines and Gödel: a reply to Mr. Lucas. *Philosophy* 37: 61, 1962) [opcional]