
知識工学 II (第 1 回)

二宮 崇 (ninomiya@cs.ehime-u.ac.jp)

この講義では、人間が持つ知識をどのように表現するのか、人間が行うような推論をどのように実現するのか、その方法について学びます。

講義内容

○記号に基づく知識表現

- ・ 命題論理
- ・ 一階述語論理
- ・ オントロジー、カテゴリ、物理的構成
- ・ 意味ネットワーク、記述論理

○確率に基づく知識表現

- ・ 完全結合確率分布
- ・ ベイジアンネットワーク(有向グラフィカルモデル)

教科書

- ・ Artificial Intelligence: A Modern Approach (3rd Edition): Stuart Russell, Peter Norvig (著), Prentice Hall, 2009
- ・ エージェントアプローチ人工知能 第2版: S.J.Russell (著), P.Norvig (著), 古川康一 (翻訳), 共立出版, 2008

知識工学のウェブサイト

<http://aiweb.cs.ehime-u.ac.jp/~ninomiya/ke2/>

はじめに (1)

この講義は知識表現を勉強することになりますが、その背景となる人工知能の歴史について初めに概説します。

人工知能はいまだ未知の領域である。

AI(=Artificial Intelligence)の定義は以下の4つに分類されます。

(1) 人間のように思考するシステム: 認知モデルアプローチ

- ・「計算機が考えるようにする、真の意味で心を持った機械を作る刺激的で新しい試み」(Haugeland, 1985)
- ・「人間の思考に関連した活動: 意志決定、問題解決、学習、…の自動化」(Bellman, 1978)

人間の思考および心のメカニズムを調査、モデル化することによりそれを再現する。

- (i) 内省
- (ii) 心理学実験

⇔ 問題の定式化による最適化。高速なアルゴリズム。

(2) 人間のように行動するシステム: チューリングテストアプローチ

- ・「人間が行う場合には知能を必要とする機能を達成する機械を作る技術」(Kurzweil, 1990)
- ・「今のところ人間のほうがうまくできている事柄を計算機にさせる研究」(Rich and Knight, 1991)

チューリングテスト: コンピュータが知能をもっているかどうか判定するテスト (Alan Turing, 1950)

人間の質問者が書面による質問を提示し、書面による返答を受け取るということを繰り返したあとで、それらの返答が人間からきたものかコンピュータからきたものか区別できなかつたら、コンピュータはテストに合格したことになる。

必要な技術: 自然言語処理、知識表現、自動推論、機械学習、視覚、ロボティクス

(3) 合理的に思考するシステム: 思考の法則によるアプローチ

- ・「計算モデルを用いた心の機能の研究」(Charniak and McDermott, 1985)

- ・「認識、推論、行為を可能にする計算の研究」(Winston, 1992)

アリストテレスの三段論法

「ソクラテスは人間である。全ての人間はいつかは死ぬ。従って、ソクラテスはいつかは死ぬ」

→論理学

- ・ 論理式で記述された解が存在する全ての問題について原理的に解くことができるプログラムが存在

問題 1: 厳密に規定されない知識や状態を論理式で記述することが難しい

問題 2: 計算量の問題 (原理的には解けるが現実の問題には事実上不可能)

合理的 ⇔ 経験的 c.f. 純粋理性批判 (Kant, 1781)

カント「純粋理性批判」入門: 黒崎政男(著), 2000 年

(4) 合理的に行動するシステム: 合理的エージェントアプローチ

- ・「知能を計算プロセスとして説明・模擬することを目的とする研究分野」(Schalkoff, 1990)
- ・「人工物の知的行動に関する研究」(Luger and Stubblefield, 1993)

エージェント: 自律動作、環境認識、長期の持続性、変化への対応、他者の目標の代行などを行う行動主体とみなせるコンピュータプログラム

合理的エージェント: 最高の結果を達成するために行動主体

合理的に思考するシステムは合理的エージェントの一部

c.f. 正しい推論 ⇔ 反射行為

AI の関連分野

哲学(紀元前 428 年～現在)

- ・アリストテレス(BC384-322)の三段論法
- ・デカルト、スピノザ、ライプニッツ: 合理主義、演繹
- ・デカルト(1596-1650): 心と物質の区別(二元論)⇔唯物論
- ・ベーコン(1561-1626)、ロック(1632-1704)、ヒューム(1711-1766): 経験主義、帰納法

- ・ カント(1724-1804): 合理主義と経験主義の統合
- ・ ヴィトゲンシュタイン (1889-1951)、ラッセル(1872-1970)、カルナップ (1891-1970):論理実証主義(20世紀前半)。全ての知識は、観察文とのつながりをもった論理的理論によって規定できる。形而上学の排除。合理主義と経験主義の統合。
- ・ カルナップ(1891-1970)とヘンペル(1905-1997): 確証性理論。如何にして経験から知識を得るか。「The logical Structure of the World」(Carnap, 1928)において経験から知識を抽出するための計算手続きを明示的に定義した。

知識と行為の関係: 目的と行為の論理的関係により正当化(アリストテレス)→2300年後に Newell と Simon による GPS プログラム。回帰プランニング。

数学 (8世紀～現在)

- ・ 正しい結論を導くための形式的な規則は何か?→論理学、アルゴリズム
- ・ 何が計算可能か?→計算不可能性、計算複雑性
- ・ 我々はどのようにして不確実な情報で推論しているのか?→確率、統計

経済学(1776年～現在)

- ・ 利得(utility)を最大にする決定を如何に行うか?→意志決定理論
- ・ ある人が他の人の活動に影響を与える「小さな経済」の場合→ゲーム理論
- ・ 利得がただちに得られず、複数の行為が連続した後に得られる場合→オペレーションズリサーチ: 数値最適化、統計学を基本にして、問題の定式化とその最適な行動を決定する

神経科学(1861年～現在)

- ・ ブローカ (1824-1880): 失語症の研究
- ・ ゴルジ (1843-1926): 神経細胞(ニューロン)の染色法
- ・ 脳活動の非侵襲計測: 1929年の Hans Berger による脳波計(electroencephalograph: EEG)、1990年の機能的磁気共鳴映像法(functional magnetic resonance imaging: fMRI)

・単純な細胞の集合で思考、行為、自意識の実現が可能である: 脳が心の座である (サール, 1992) ⇔ 神秘主義的な考え方

しかし、まだ認知活動の実現の仕組みを理解するにはまだまだ遠い道のりがある。

スパコン: 1 万 CPU(10^4), 1 兆トランジスタ(10^{12}) 1GHz(10^9)

PC: 4CPU, 10 億トランジスタ(10^9), 1GHz(10^9)

人間: 1000 億ニューロン(10^{11}), 1MHz(10^3)

心理学(1879 年～現在)

- ・ヘルムホルツ (1821-1894)とヴント(1832-1920): 視覚の研究に科学的方法論。
- ・行動主義: 内省が得られない動物に対する客観的実験方法を人間に適用する。刺激→動物→反応。ネズミや鳩には多くの発見があったが、人間に対してはあまり成功しなかった。
- ・認知心理学、認知科学: 脳が情報処理装置であるとして、人間の認知活動を明らかにする。知覚には意識下の論理的推論が必要である。信念や目的などの心的用語。刺激→内的表象→認知プロセスによる新しい内的表象→行為

コンピュータ技術(1940 年～現在)

効率のよいコンピュータをどのようにして作るか?

制御理論とサイバネティクス(1948 年～現在)

- ・人工物はどのようにして自分自身の制御下で動作できるか?
- ・制御理論: 目的をもった行動を、現状と目的状態の差分である「誤差」を最小化しようとする調整機構とみなす。目的関数を時間と共に最大化するシステムの設計。
- ・制御理論と AI: 制御理論「微積分・行列による固定された変数で記述できるシステム」⇔AI「制御理論数学の限界を越えた研究」

言語学(1957年～現在)

- ・ 言語は思考とどのような関係にあるのか？→自然言語処理、知識表現
- ・ チョムスキー: 行動主義、構造言語学の批判。文法理論の定式化。生成文法。

AIの歴史

人工知能の懐胎 (1943～1955): ニューロンのモデル化、チューリングテスト、遺伝的アルゴリズム、強化学習

人工知能の誕生 (1956): ダートマスワークショップ。創造性、自己改善、言語使用といった人間の機能の複製を目標に。

初期の情熱、大きな期待 (1952～1969): 一般問題解決器 (GPS)。幾何定理証明器。強いアマチュア程度に強いチェッカーゲームの学習プログラム。高水準プログラミング言語 Lisp。Advice Taker: 実世界の知識表現と推論(一般知識)をもった問題解決器。知能テストにでてくるような問題を解くプログラム。パーセプトロンとその収束定理。

現実からの反撃 (1966～1974): より大きく広範囲な現実的な問題に対しては無力であった。計算複雑性。ニューラルネットの学習性能に対する誤解。

知識に基づくシステム: カへの鍵となるか? (1969～1979): 専門のタスクに特化したエキスパートシステム。血液感染症診断システムは専門家並の診断成績。知識表現や推論言語の開発。Minskyのフレーム。

AIが産業となる (1980～現在): 計算機システム注文の助言を行うエキスパートシステムにより年間数千万ドルの経費削減。全体として数十億ドルの売り上げ。

ニューラルネットの復活(1986～現在): ニューラルネットワークのより深い研究。コネクショニズム
⇔記号的アプローチ

AIが科学となる (1987～現在): 経験的実験による検証。HMM, データマイニング、ベイジアンネット。

知的エージェントの出現(1995～現在): マルチエージェント、Web技術

超大規模データセットの利用 (2001～現在): 語義曖昧性解消の教師無し学習。知識のボトルネックの解消。ハンドコーディングによる知識工学でなくとも、十分なデータがあれば学習により解決。