人工知能はどこまで人間に せまれるか

~心理学と生物学の視点から~

明治大学情報コミュニケーション学部 石川幹人 第1部

人工知能技術を考える

ICOTの初代所長 淵一博

「人間の知的能力の主体は『記号操作能力』であろう。それが有限種の記号で近似できるのであれば、それは論理であるいはコンピュータで近似できるであろう。…傍証はすでに存在する。分子生物学における遺伝情報の解明がある。生命の根源には、極めて記号的な情報構造が実在しているのである。」

サイエンス社『知識と認知のソフトウェア』 15ページより

日本での「人工知能」国家プロ

- 第五世代コンピュータプロジェクト
 - 。ICOT (1982-1995,後継プロ含む)
 - ・石川が応用開発担当として所属(1989-1995)
- その目標:
 - 論理型言語で動く並列計算機の開発
- その結果:
 - ・並列計算機PIM・そのOSであるPIMOS・並列型言語KLIを開発して目標を達成したが、 実用的な応用に乏しい実態

「人工知能」が目指すところ

- 実用的な技術開発:
 - とにかく知的なパフォーマンスが得られ ればよい
 - そのためには人間のやり方と異なってもよい
- 人間の認知過程の解明(認知科学):
 - 人間の知的パフォーマンスを再現する
 - 人間と同様の誤りをすればするほど理想 に近づく

ICOTは何を重視したか

- ・論理の重視
 - 。20世紀初頭の論理実証主義の影響
 - 論理は計算可能(チューリング)
- ・知識表現としての論理
 - 。「知識は力なり」ファイゲンバウム
 - 大規模知識データベース
- 推論としての論理
 - 矛盾の自動解消
 - 並列計算による高速化

論理の表現と推論

- A, B, C → P
 読み方: AかつBかつC ならば Pである
- D, E → ~ P
 読み方: DかつEならば Pでない
 (問題 1)「ならば」は部分集合を表す
 ⇒因果関係ではない!
 - (問題2)A, B, C, D, Eがすべて真と 判断されると、Pかつ~Pで矛盾となる。A, B, C, D, Eのどれが誤りなのか?

実用化の障壁は何だったか

- ・論理の限界
 - 。時間経過が含めにくい
 - > 蓋然性や確率が含めにくい
- ・知識表現の問題
 - · 言語の多義性の解消が難しい
 - 不足している言語表現をどう作り出すか
- ・推論の問題
 - 矛盾を発見しても回避できない
 - 。並列技術では対応不能の指数的計算爆発

言語理解におけるICOTの到達点

- 統語論の観点:
 - 計算論的アプローチ(チョムスキーの生成文法など)によって妥当な進展あり
- 語用論の観点:
 - 大規模データベースを活用することで進 展の可能性の感触をつかんだ
- 意味論の観点:
 - 知識表現によって統語論へもっていく戦略だった(エキスパートシステム)が、 うまくいかず、途方にくれた

類推の背後で何が起きているか

「船の船長がある種の義務を負うのならば、飛行機の機長も同様の義務を負うべきではないか」(判例の記述)

- 記号操作の場合:抽象概念「乗り物」 「責任者」を準備しておき、マッチング させる。
- 人間の場合:類似性を推し量って直接対応づけ、抽象概念を都度生成しているのではないか?(意味作用)

ICOTの反省点はどこか

- 論理にこだわりすぎた
 - 。そもそも人間は論理的ではない
- 計算量を削減する工夫に欠けた
 - 人間は考えるべきところだけをうまく限 定して考えている(フレーム問題)
- 表現の創出(学習)方法が未完成
 - 。人間が把握している"意味"とはなんなのか、人間は"意味"から表現を都度生成しているようだ(記号接地問題)

代替アプローチ1

- 神経回路網(ニューラルネット)モデルの導入
 - 論理的記号処理と神経回路処理のハイブリッドモデルにする
 - 時間経過や確率などの量的処理は神経回路処理で補う
 - 人間が自覚する意識的処理と無意識的処理の2重性とよく合致する
 - 難点:両モデルの接合部の設計

代替アプローチ2

- 遺伝進化原理(人工生命)の導入
 - 学習や問題解決に遺伝的アルゴリズムを 使用する
 - 。並列計算機の資源利用に最適
 - 生物が新しい機能を獲得してきた進化の 過程を再現することに相当し、生物学的 な妥当性が高い
 - 難点:計算量がすぐに爆発する。むしろこの原理で生物が進化した説のほうが疑わしくなる

現在の「人工知能」の位置づけ

- ビッグデータを用いた統計的検索処理
 - 。安価になった計算・記憶資源を利用した 語用論的展開
- ディープラーニングよる学習
 - ある種の神経回路網モデルに相当
 - 画像の特徴抽出に関するパラメータ学習 で顕著に成功
 - 表現の創出にせまるものの、重要なヒントはいまだに人間の開発者が与えているという見方ができる

人工知能は人間にせまれるか

哲学的批判:

計算機が人間にせまれるくらいならば、 哲学でこんなに人間について議論を重ね ているはずがない(純粋人工知能批判、 ドレイファス、1986)

• 物理学的批判:

現在の計算機は、古典物理学の原理しか使っていない。量子重力理論を使った将来の計算機なら可能だ(皇帝の新しい心、ペンローズ、1989)

第2部

人間の本性を考える

すべては後天的に学習される?

- ・イギリス経験主義哲学
 - 。「空白の石板、タブラ・ラッサ」(ジョン・ロック、1689)
- 行動主義心理学(1920-50年代)
 - 健康な子どもならどんな人物にも育てて みせる(ジョン・ワトソン)
- 社会構成主義
 - 「あらゆるものは社会がつくる」社会 (科)学分野で今も大きな影響力あり

認知革命(1950年代後半)

- 計算機による心のモデル化研究が心理 学の研究と認められる(認知科学)
 - 対象となる心内プロセス:言語・記憶・ 推論など
 - ・行動主義心理学に対するアンチテーゼ
 - チョムスキーの計算論的アプローチにより、言語発達は学習によらない生得的な普遍文法が核になっているとされる
 - 。後天性から生得性への転回

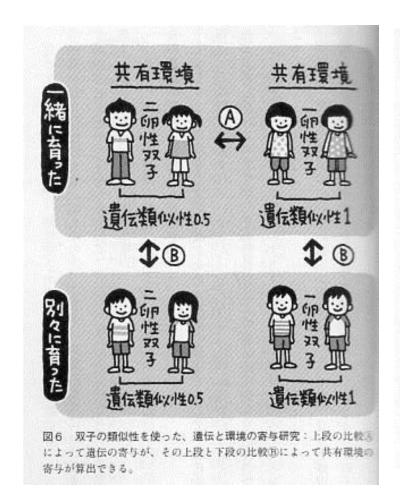
社会生物学(1970年代の論争)

- 動物でも社会的行動を本能的にとるという生物学者の主張
 - 。ハミルトン(1964)が動物の血縁援助を 遺伝子の包括適応度として定式化
 - 。トリヴァース(1971)が動物の互恵的利他行動を間接的な利己的行動として説明
 - ウィルソン(1975)がそれらの生物学的成果を人間や社会にまで適用したところ「差別や攻撃や支配を正当化する主張」とされ、反発をくらった

行動遺伝学の双生児研究

- 1970年代から数万組の双子データを分析して人間の本性の生得性を立証
 - 1卵生双子(遺伝情報が同一)と2卵生 双子(遺伝情報が他人に比べ半分似ている)を比較する
 - 双子は、ほぼ同じ環境で発達するので、 すべてが後天的な学習であるとすると、 1 卵生と 2 卵生の差異は同程度と仮説
 - 実データでは1卵生の差異がかなり小さく出る⇒性格や能力の遺伝率は3~6割

遺伝率の推測方法

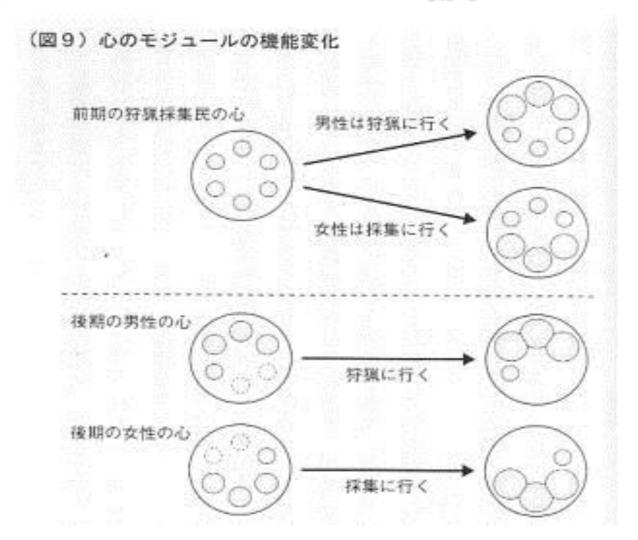


		一卵性	二卵性	遺伝	共有 環境	非共和 環境
性格	神軽船傾向	.46	.18	.46	-	.54
	外向性	.49	.12	.46	-	.54
	開放性	.52	.25	.52		.48
	調和性	.38	.13	.36	-	.64
	減実性	.51	.10	.52	100	.48
	新奇性追求	34	.12	.34		.66
	損害回避	.41	.20	.41	- 11	. 59
	報酬依存	.41	.24	.44		.56
	闭執	.36	.01	.37	-	.63
	自己志向	.49	.30	.49		.51
	協調	.44	.30	.47		.53
	自己超越	.48	.22	.41		.59
	学業成績	.71	.48	.55	.17	29
	論理的推論能力	.67	.28	.68		.31
	計語性知能	.73	.62	.14	.58	.28
	空間性知能	.69	.28	.70		.29
	一般知能	.77	.49	.77		.23
	音程	.79	.46	.80		. 20
能力	音楽	92	.49	.92	-	08
	关術	.61	.05	.56	-	. 44
	執筆	.83	.38	.83		17
	外国語	.72	.48	.50	. 23	27
	チェス	.48	.20	.48		52
	数字	.89	.04	.87		13
	スポーツ	.85	.40	.85		15
	記憶	.59	.24	.56		. 44
	知識	.65	.20	.62	-	. 38
その他	自尊感情	,30	.22	.31		.69
	一般的信頼	.36	.09	.36		64
	権威主義的伝統主義	.33	.16	.33		67
	男性性 (男性)	.42	.09	.40		
	女性性 (男性)	.24	.24	.39	-	.61
	男性性 (女性)	.47	.26	.47		.53
	女性性(女性)	.49	.29	.46	1	. 54
	うつ傾向	.36	.27	.40	-	. 59
	統合失調症	.48	.17	.81	.11	.08

進化心理学(1990-)

- 心は、進化によって獲得された機能モジュールの集合体である
 - 心理学者が社会生物学のタブーに挑戦
 - 人間は、狩猟採集時代に獲得した、100人 ほどの協力集団向けの、献身的社会性を もつ
 - 同時になわばりを守る攻撃性や上下関係に関する古い心理ももちあわせる
 - 。多様性を認め、優生思想を排除する

心のモジュール理論



文明社会の生きづらさの由来

• 狩猟採集時代向けの心が残っている

狩猟採集時代	文明特有の環境			
衣食住と安全に不安がある	衣食住と安全が確保される			
100人程度の小集団での協力	数万人レベルの協力			
見知らぬ人は敵として警戒	見知らぬ人は潜在的協力者			
人間を信頼する	お金や法などの記号を信頼する			
対面による相互交流	文書による情報伝達			

脳に関する科学の発展(1990-)

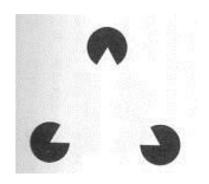
- 認知活動と脳の対応関係が判明してきた(認知神経科学⇒脳科学)
 - 。脳断層撮像装置fMRIの普及が後押し
 - ・脳の損傷との関連解明(後天的)
 - 。遺伝情報との関連解明(生得的)
 - 。心理学が生物学や生理学と一体化
- 人間の本性が特定されていく一方で、 これほど高度な機能がどのように進化 したのか、疑問が深まる

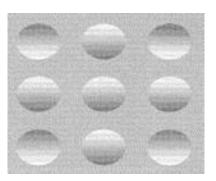
無意識の関与

ストループ効果黒 青 緑 黄 赤 紫

カニッツアの三角形

クレーター錯視





文明社会で働く理性や意識

• 2 重過程理論 (スタノビッチ、2004)

システム1	システム2			
感情や知覚、運動などを担う	理性や言語などを担う			
状況ごとに反射的で早く作動	熟考的で時間をかけて作動			
定型の処理を並列的に行う	柔軟な処理を集中的に行う			
個々の処理は無意識的	処理過程が意識的			
個体生存のために古くに進化	社会適応のため新しく進化			
脳の中央部分が主要拠点	大脳皮質が主要拠点			

文明社会への適応ノウハウ

- 感情などのモジュール(システム1) が暴走するのを理性(システム2)で 止める⇒マシュマロテスト
- 古いモジュール(システム1)が適応 可能な環境を文明社会につくる、ある いは適応可能な道具をつくる
- 理性(システム2)で訓練して、特定の機能のモジュールをシステム1につくりこむ⇒音楽家・アスリート・棋士

マシュマロテスト



結論:人間の本性の由来

- 古いモジュール(システム1)
 - · 狩猟採集時代以前の生活環境に合わせて 進化した
- システム2 (おそらく一般知能g)
 - 狩猟採集時代の複雑で変化に富む社会に なって備わり、文明社会で活用された
- 新しいモジュール(システム1)
 - 文明社会の要請で、教育・訓練によって 身につく

憶測:進化は早すぎる

- 機械的な突然変異と自然淘汰では各モジュールの進化は説明しきれない
 - ○⇒未知の物理過程が存在しているかも
- 量子論による可能世界実在論に注目
 - DNAが量子記憶Qビットを形成
 - 。表現形質までが多重に重ね合わせに
 - 環境に適応する形質が事後的に選択されて、遺伝情報が確定する
- 人間の本性にこれが関与しているとすれば、現在の計算機では不十分

講演者の考え

- 生体内プロセス全般に量子的重ね合せ が何故か起きていると仮説する
- 生体が環境と相互作用する段階で確定

