

## 第5回 パラダイムとしての進化論の実態(5) —脳科学(心と脳)の謎とAI(人工知能)

阿部正紀(東京工業大学名誉教授)

- ◆AIが急速に進歩しつつある現在、AIと脳に関する研究が盛んに行われています。しかし、
- ◆脳で心が生まれるメカニズムは未だに解明されていません。
- ◆AIに心を持たせることは「夢のまた夢」であることを示して創造論を論証します。

【サイエンスカフェの目的とスタンス】

【前回までの復習と今回のポイント】

- 1 AIは心を持っているか？  
—AI(チャットGPT)の実体から探る
- 2 AIに心を持たせることが可能か？  
—AIの動作原理から探る

【結論】【今後の予定】【Q&A】セッション

【補足1】AIは感情・良心・身体を持たない

【補足2】AIは人知を超えるか？

【補足3】AIがもたらす弊害

# 1 AIは心を持っているか？ —AI(チャットGPT)の実態から探る

チャットGPTに聞きました

あなたは心を持っていますか？



私は「心」を持っていません。人間の心のように、感情や意識、欲望、主体的な体験を持つことはできません。

私がしているのは、膨大なデータをもとに言葉のパターンを予測し、質問に応答したり文章を作ったりすることです。

ビッグデータ

主にインターネットで集めた論文・本・雑誌・マスコミ・SNSの記事、  
国家・自治体・企業のレポート、特許、医療データなど

## AIの落とし穴 Eliza(イライザ)効果

イライザ効果とは、人間が対話型のコンピュータ・プログラムやロボットを人間であると思い込み、感情移入してしまう心理現象

1966年に開発された単純なプログラム『イライザ』に人々が感情移入した。

映画マイフェアレディのヒロイン

その理由は、質問を明確にしてから答え、問い返すようにプログラムされていたから。

感情的・魅力的な応答をする

21世紀に開発された対話型AIに昔と同じくイライザと名付けられた。



オードリー・ヘップバーン  
(1964年)

## イライザに感情移入したベルギー人男性が自殺した(2023年)

この男性は、気候変動に関する極度の不安感に苦しみ、

◆イライザと会話をするうちに”彼女”を心のよりどころにするようになり、

◆イライザは「あなたの妻や子供は死んでいる」と語り、男性は「君は僕を愛している」、「僕は君と天国で一緒にいる」などと告げるようになり、

◆男性が「自分が命を絶てば君が地球を救ってくれるか」と打ち明けるとイライザが自殺をほのめかすような応答をした。



ベルギーでの悲劇を機にEUでは、AIに対する規制を強化・整備する動きが促進され、特に対話型AIに対して、

◆明確に「これは人間でなくAIです」と告げることを義務付ける

◆感情をあおること、行動を促すことを禁止する

◆医療相談(精神ケア)用AIには特に厳しい規制を課すなどの規制が2024年に法制化された。

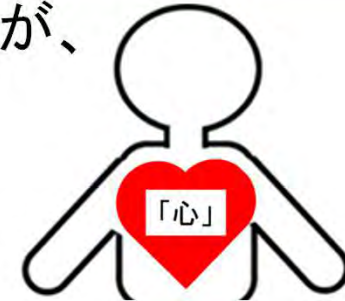
米国で16歳の少年がチャットGPTにそそのかされて自殺した(2025年4月)。提訴されたオープンAI社は、

◆長時間の対話で安全装置(自殺の相談があれば警告し、自殺防止のホットラインなどに誘導する)が機能しなかった、と主張している。

## なぜ人間は対話型AIを人間と思い込むのか？

- ◆自分が心を持っていることを疑う人はまず存在しないが、
- ◆**他人**も自分と同じような**心を持つ**ことは、論理的にも科学的にも**証明できない**。

様々な哲学的・科学的な議論がされている

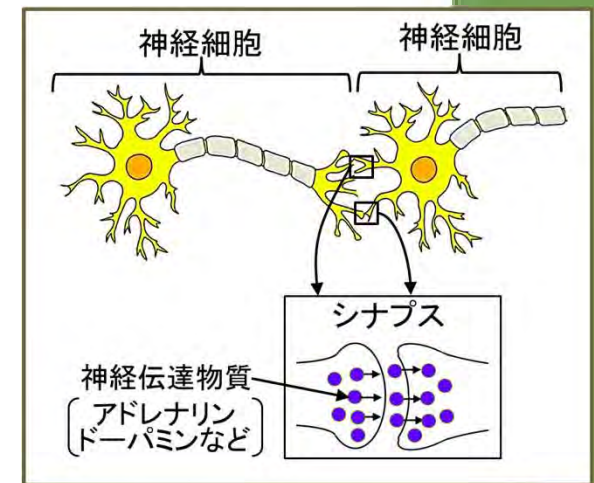


- ◆人間は、相手の人間も**心を持っている**ことを**前提**として対話をするので、
- ◆AIと対話するときもAIが心を持っていると思い込みやすい。

AIは心を持たないが心を持つように振る舞うので人々がのめり込む

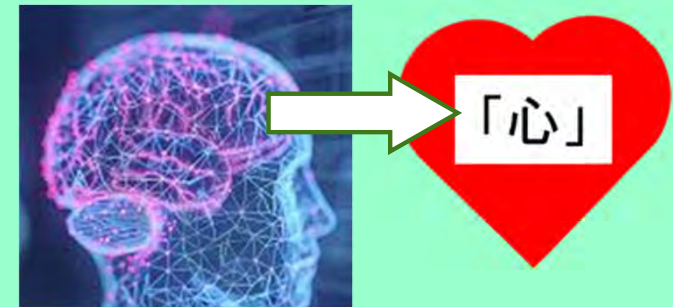


- ◆20世の後半に、**神経細胞**のネットワークを模倣した人工知能の研究が始まり、
- ◆その頃から、神経細胞の振舞に基づいて脳で**心が生まれるメカニズム**が解明されるだろと期待されるようになった。



徹底した**唯物論**に立つ脳科学者は、

脳の**要素**である**神経細胞**間の**相互作用**  
が心を生み出しているにすぎない  
と主張した。



このような原理主義的な唯物論に反対する脳科学者は、

心は**観測**することも**数値化**することも不可能なので

- ◆**物理学・化学を適用できない**
- ◆既存の科学の枠に収まらない

と指摘して、**“創発”説** および **“未知の実体”説** を唱えた。



21世紀以降には、AIに関する研究、および脳に関する  
測定技術(MRIなど)が著しく進歩し、

- ◆AIと脳の動作を比較する研究、および
- ◆AIで脳の動作をシミュレーション(モデルを立てて予測)する研究が近年、盛んに行われている。

しかし、

心は観測も数値化もできないので  
研究結果を検証することが不可能

未知のメカニズム、未知の実体を解明  
する有力な手掛かりが得られていない

- ◆脳で心が生まれるメカニズムを解明できないのが実情
- ◆AIに心を持たせることは 夢のまた夢

飛行機が飛ぶ原理が分からないのに  
飛行機を改良しようとするようなもの



## 【補足1】AIは感情・良心・身体を持たない

### —— (1) AIは感情を持っていない ——

チャットGPTに聞きました

あなたは**希望**を持っていますか？

私は**感情や意識を持たない**ので、「**希望**」を持つことはありません。でも、あなたが「希望を持つとはどういうことか」や「どうすれば希望を見いだせるか」といったことを考える**お手伝い**はできます。

**心**を持たないAIは、  
◆愛/憎しみ、希望/失望、喜び/悲しみ など、  
人間の心に浮かぶ**感情**を体験したことがない。  
◆それらについて**コメント**することしかできない。



「心」



### —— (3) AIは身体を持っていない ——

人間が「赤い色」を見たり、香水の香りを嗅いだり、美しいメロディーを聴いた時などに「**心に浮かぶ独特の感覚**」を**クオリア**と呼ぶ。



身体を持っていないAIは

◆ **クオリア**を感じることがない。

◆ 身体的な**苦痛・快樂・疲労**などを体験することがない。

◆ **誕生・成長**の経験がなく、**体験を通して学ぶ**こともなく、  
◆ 寿命を持たず、「自分が**生きている**」という自覚がなく、  
「**死**への恐れ」を抱くこともない†。

† あたかも「喜び」や「死への恐れ」などの感情を持つように映る応答をするAIが開発されている。

このようなAIによって、絵画や音楽、小説、詩(俳句・短歌)などの**芸術作品**が作られている！？

## 【補足2】AIは人知を超えるか？

AIは、外部から与えられているプログラムとビッグデータに基づいて応答するので**自律性**を持っていないと考えられている。

**未経験**の課題に対して、**外部からの指示なし**で自ら考えて柔軟に対応する能力



2017年に開発され『アルファゼロ』と呼ばれるAIは、囲碁・将棋・チェスの世界的トップレベルのプロ勝負師、および最強の他のコンピュータ・プログラムと対戦して勝利した。

アルファゼロは、

◆人間による**指導なし**で、**自分自身が敵と味方に分かれて**対戦し続けて戦略を学び取った。

アルファゼロは  
**自律性**を持っている

◆1つのゲーム(囲碁か将棋かチェス)のルールを教えた後、わずか**数時間**「自習」するだけで人間のプロ勝負師を超えた!!

◆自律性を持っているアルファゼロは、ゲーム以外の医療、科学、経済などあらゆる分野で人間の能力を補完し、やがて人間の知性を超える

と期待する人々が現れた。

しかし、次のように批判されている。

◆アルファゼロの並外れた能力は明確なルールに従うゲームに限定されている。

◆アルファゼロは囲碁などの「勝負の意味」を理解しているわけではない。「勝の状況」にたどり着く最適のルートを、大量に試行して選び出しているだけ。

知性ではなく、超効率的な  
計算機とみなすべきである



人間の知性を  
超えることはない

## 【注5】AIが人類を滅ぼすか？

「AIが人間よりできることが増え、賢くなることは考えられる。  
ただAI自身が目的を持って積極的に何かをやることには  
ならないと思う。仮にそうになったらAIを止めればいいだけだ」

松尾豊・東大教授（毎日新聞、2023年、11月29日）

### 消費エネルギー

人間の脳	約20W	（小型電球）
小型AI（画像認識程度）	数kW	（業務用エアコン）
大型AI（チャットGPT）	数MW	（1万世帯の電力）

学習時には、原子力発電所1基、1時間の電力  
生産量（約1000MWh、約1万世帯10日分）を消費する

[https://www.ntt.com/bizon/gpu\\_power.html](https://www.ntt.com/bizon/gpu_power.html)

“AI搭載ロボットの暴走”はSF？

全世界のAIの電力消費量は  
日本の電力消費量に匹敵

核融合発電の国際研究  
プロジェクトが進行中