

# 思考の「ブラックボックス」

## 『人間の未来 AI の未来』を読んで

上海交通大学

医学院 医学学部 1年

呉頤沢

「AI にバッハの曲は作れても春樹の小説は書けない」

この本の中で、iPS 細胞研究所の山中伸弥所長と将棋の羽生善治棋士が対談し、この結論に至っています。

起源を遡ると、AI は人工の「知能」、人間は自然の「知能」で、人工とは自然を模して作り出すことです。両者に何の相違があるかを討論するとき避けて通れない問題は、人間の持つ「高い知能」の知能は一体何かということです。

しかしそれはブラックボックスで、現時点では、一体何が起きて人間が際立っている高い知能を持つことになったのか見いだすことはできません。AI は統計とデータに頼って、結果から逆算して人間の思考パターンを近似することしかできず、その経緯を真に復元することはできません。

これを踏まえると、AI は最適化、つまり組み合わせから最も適切な答えを見つけることが得意です。AI の各領域における進展の本質はアルゴリズム、データの蓄積、データの処理能力と関係しています。

「AI はバッハの音楽を作曲できる」という評価は AI の模倣能力の肯定です。バッハの音楽の二重盲検テストにより、このことが確認されました。「本物のバッハ作品」「バッハ風の AI の作品」「バッハ風の人間の作品」の3種類の音楽の中で、ほとんどの人が最も機械的だと感じた音楽は、実は「バッハ風に人間が作った」音楽で、「バッハ風に AI が作った」音楽ではありませんでした。

この事例から、人間と AI の観点からそれぞれの特徴が反映されます。人間の模倣は主観的なものを含み、模倣する者と模倣される者の混合体で、千人の目には千人のハムレットがいます。「バッハ風に

人間が作った」音楽は、最も機械的だと感じたとされていますが、本物のバッハ作品の作風は備えており、バッハの作風に対する人々の異なる理解の表れです。

そして、AI の模倣はより同一になる傾向があります。同じ作風は、複雑に見えるだけで実際には表面的な固定フォーマットにすぎません。上記のテストにおいて、AI が人間よりも優れている理由は、評価基準に対して与えられた答えを出し、データを処理することで、一般的な意味での機械的な感覚が弱まり、「バッハ風」の復元を強化できるからです。

結果から、AI の強みはその膨大なデータベースにあることが容易にわかります。これにより、「量」の面で大きな強みが生まれます。対照的に、人間が頼りにするデータベースは比較的限られているため、経験の限界により、意思決定の際に十分な比較オプションが不足することになります。

また、AI の客観的な分析能力は、人間の本能的、主観的な欠陥を適切に補完します。人間が個性的な美を追い求めるからといって、AI が結果を出すためのあらゆる試みを放棄することはありません。「既成概念」の誤解に陥ることもありません。いかなる可能性も自発的に排除せず、比較して最善の選択肢を見つけるのです。

この点で、模倣という行為は、実際には既存の経験をまとめたものです。人間の過剰な主観的感情は、時によっては客観的な AI の創造物ほど柔軟ではなく、保守的で進歩しないのかもしれませんが。

そのため、人間の観点から AI のもたらすものとの共存を考えると、実用上の効率性と利便性に加えて、さらに重要なのは、先入観を捨ててこそ他人の長所を広く吸収することができるという啓発です。将来、人間は AI の巨大なデータベースを利用して、客観的な選択肢を広げ、「ブラックボックス」の過度に主観的な制約から脱却し、新しい知識をオープンに受け入れられるようになります。

しかし、AI は思考のブラックボックスにおける主観の弱点を放棄したと同時に、人間の「第二の天性」、つまり超越する能力も失いました。

「村上春樹の小説は書けない」のは、小説に含まれる感情を AI が理解し表現することが困難だから

らです。二重盲検テストで検討されなかった「本物のバッハ作品」の中には、AI では再現できないバッハの感情の深みが隠されています。これらの作品は、過去の経験を単に統合したものではなく、人が自分自身を積極的に探求するときに考えたり感じたりすることを表現したものです。

このプロセスも「ブラックボックス」を通りますが、完全に「直感」を使って既存の経験を超越しており、もはや順列と組み合わせの最適解ではなくなり、結果に対する評価基準もありません。ひらめきと感情の昇華なのです。

進化の観点から、生物が目を進化させるためには、他の器官が鈍くなるため、いわゆる「直感」も、進化の過程で鈍ってしまった機能を再び活性化させるためなのかもしれないという考えがこの本に出てきます。日本語には「直感」、「勘」という言葉がありますが、前者は先天的な能力、後者は後天的な経験を指すときに多く使われます。

これは、ある一定の「量」の蓄積によって無意識のうちに得られる「質」の変化です。そこには、私たちが「第二の天性」と呼ぶ本能、つまり既存の「量」を超越する潜在能力が存在します。そして、この潜在能力こそが、人間が持つ真の「知能」だと思うのです。

AI がイベントを処理する手順と比較すると、データへの過度の依存により、ある程度、既存の経験を抽出する能力も失われています。その結果、限界を超越することができず、既存の経験に縛られることしかできなくなります。この点で、まさに AI の限界が反映されています。

要約すると、AI は事実の判断に基づいてのみ最適化することができ、これは「量」の範囲内での比較であり、多くの場合、厳格な行動パターンを使用します。しかし、人間の超越する本能は「量」を突破し「質」を変えることができるのです。価値判断に基づく主観、理性と感情を織り交ぜたものであり、そして人間としての昇華の具現化です。そのため、「直感」の背後にある第二の天性は、これまでのところ、人間に特有の特徴であり、人工「知能」が模倣したり比較したりできないものです。

だからこそ羽生氏は「やっぱり個々の人間が肉体を持って生活してきた背景があって書くからこそ、作

品として意味を持つ」と言ったのでしょう。これは「質」の変化における人間の超越性の肯定です。

ここから、AI の角度から共存共栄を考える鍵となるのは、「直感」の本質をどこまで探求できるか、そして AI がどこまで同じ機能を実現できるかという点だと分かります。「ブラックボックス」の中にある直感の科学的性質を発見しない限り、AI が「量」から「質」への変化を実現できることはありません。

現在、AI と人間の活動の原理に基づいて、両者の境界を明確に定義することができます。つまり、人間が主観によって干渉された自然の本性を失うことは不可能で、AI も人間の超越という第二の天性を学習できないこと、羽生氏の対局における明言しがたい第六感、正義のために命を捨てる信念です。

そして、AI のもたらすものとの共存を考えることは、「知能」について考えることであり、最終的に本人に戻ってきて、思考の「ブラックボックス」を開くことです。

山中伸弥『人類的未来、AI 的未来』上海訳文出版社、2022. Print. 訳文視野  
(原著は山中伸弥、羽生善治著『人間の未来 AI の未来』)