

[人類はどう生きるべきか？ ITはどうあるべきか？]

<人間とは何か>を 超越した知性

応
般

7.4

黒崎政男 (東京女子大学)

かつて、D. Hofstadter は『ゲーデル・エッシャー・バッハ』(1979)^{☆1}のなかで、次のように述べていた。
「人工知能とは、なんであれ、まだ、なされていないところのものである」

つまり、こういうことだ。人工知能の進歩に関しては、1つの定理とも言うべきものがあって、何らかの心的機能がいったんプログラムされると、人々はそれを「真の思考」の本質的部分と見なすことをやめてしまう。知識の不可欠の核心は常に、次のいまだプログラムされていない事柄の中にある、というわけである。

このHofstadterの定理に従うと、人工知能は永遠に実現しないことになる。まだ、作られたことのないものが人工知能だというのが、人工知能の定義だとすれば、原理的に、人工知能など作れるわけがない。作ったとたん、つまり、神秘的と思われていた人間のある能力の仕組みが分かり、それを定式化し得たとたん、その能力は知能とは呼ばれなくなるのだから。

人工知能のゲーム研究の分野で、9つの升目を使う「三目並べ」で、コンピュータが常に勝つか引き分けるかを可能にしたとき、反AI論者のH. L. Dreyfusは次のように言った^{☆2}。

「コンピュータは単純な三目並べはできるが、本当の知能が必要なチェスは指すことができない」

1997年チェス・コンピュータ、ディープ・ブルーが、人間のチェス世界チャンピオンのG. K. Kasparovを破ったとき、私は、次のように考えた^{☆3}。

「1秒間に数億手考えるコンピュータと、数手浮かんだだけでそれに匹敵する人間の直観力や洞察力はすばらしい。無意味ほど膨大なしらみつぶし計算をしなければならないコンピュータのやり方

はほんとうの知能とは言えない」

だから、より高度な直観や洞察を必要とする将棋は、コンピュータには指せないだろう、と思われたわけだ。2010年代に入って、コンピュータの指す将棋が、次々とプロ棋士を破り始めたとき、我々はどうのように語るべきなのだろうか。

「将棋を指せるぐらいでは、本当の知能とはいえない。ほんとうに全体の洞察力が必要な囲碁を、コンピュータは打つことはできないだろう」

だが、早晩、コンピュータ囲碁も人間の囲碁名人に勝つ日がやってくるだろう、と考えるようになった我々は次のような言葉を考えるのだろうか。

「囲碁が打てるぐらいでは本当の知能とは言えない。結局、ゲームというのは厳密な規則に従っているだけだからだ。自然言語を楽々と理解するようであれば、本当の知能とは言えない」

そうこうしているあいだに、IBMが開発した質問応答システムの人工知能「Watson」は、2011年、伝統的なクイズ番組に登場して、自然言語をいわば「理解」し、的確な解答をして、人間チャンピオンを打ち破ってしまった。

三目並べはできるが、チェスはできない。チェスはできるが、将棋ができない。将棋はできるが、囲碁はできない。囲碁はできるが言語は理解できない。言語は理解できるが、**はできない。このような形で、そのつど要求水準を上げていくこと

^{☆1} Hofstadter, D. R.: Gödel, Escher, Bach: An Eternal Golden Braid (1979), ダグラス・ホフスタッター 著、野崎昭弘、はやしはじめ、柳瀬尚紀 訳:ゲーデル、エッシャー、バッハーあるいは不思議の環、白揚社 (1985)。

^{☆2} Dreyfus, H. L.: What Computers Can't Do - The Limits of Artificial Intelligence - (1972/79), ヒューバート・L. ドレイファス、黒崎政男、村若 修 訳: コンピュータには何ができないか? 哲学的人工知能批判、産業図書 (1992)。

^{☆3} 黒崎政男: となりのアンドロイド、NHK出版 (1998)。

で（Hofstadterの定理によれば）人工知能はいつでも不可能ということになってきた。だが逆に言えば、人工知能は、この数十年でどんどん進化してきた、ということだ。

Kurzweilは「これから数十年のうち」に、「情報テクノロジーが、人間の知識や技量をすべて包含」することになり、「ついには人間の脳に備わった、パターン認識力や、問題解決能力」、さらには「感情や道徳にかかわる知能すらも取り込むようになる」と論じている^{☆4}。

それが2045年だというのが、私には、それがそんなに遠い未来の話ではないように思われる。その日は「やって来る」、というよりは、日々、すでにその日に「なって」いつているように思われる。

ただし、それはこれまでの人間知能が解明されたうえで、こうした変化が起こるのではないだろう。従来、人工知能研究のおもしろさは、コンピュータにおける知能の限界を知ることによって、逆に人間知能の特質や広がりについて何かを学ぶことができるという点にあった。コンピュータによる知能のシミュレーション、その限界を通して逆照射される人間知能の特質を知ることがAI研究の意味でもあった。

だが、たとえば、Watsonは、総メモリ容量15TB、総プロセッサ・コア数は2,880個で構成されており、約100万冊分の書籍をデータベースとして持っているという。人間の知能、人間のクイズチャンピオンがたどる思考法とは、まったく異なる

種類のテクノロジーで、人間の知能以上のことを実現している。

さらに言い換えてみれば、こういうことだ。〈飛ぶ〉という働きを人工的に実現するため、鳥の羽や筋肉の構造や動きを理解しシミュレートしようとする努力とは別に、ジェットエンジンというまったく異なる種類のテクノロジーで、飛ぶことを実現する、しかも、遥かに効率的な形で。

こうなると、人工知能の進化を追うことで、人間の知性とは何かに対する新たな知見がもたらされることを期待するという立場での人工知能研究は、当初の意味を失ってゆくだろう。人間みずからの自己理解を模倣し、後追いすることから生まれた人工知能は、「人間ならできる、こんなことができない、まだ、あれができない」と言われているうちに、いつのまにか独自の進化を遂げてしまう。その自律的発展はまったく別のルートを取りつつ、人間の知性を超えるという目的そのものを達成しつつある。そこが人工知能の、いわば人類の予想を超えた「跳躍」であり、〈超知性〉たるゆえんなのだろう。

(2014年10月8日受付)

黒崎政男 | PFH00600@nifty.com

1954年仙台市生まれ東京大学文学部哲学科卒業。同大学院博士課程修了。現在、東京女子大学現代教養学部人文学科哲学専攻教授。専門は、カント哲学、人工知能論。著書に『哲学者はアンドロイドの夢を見たか』など多数。

^{☆4} Kurzweil, R.: The Singularity Is Near: When Humans Transcend Biology (2005) (レイ・カーツワイル著、井上 健監訳他:ポスト・ヒューマン誕生—コンピュータが人類の知性を超えるとき—, NHK出版 (2007).