

# AIの来し方ゆく末

西垣 通 東京経済大学コミュニケーション学部教授 東京大学名誉教授

## ◆ 思考機械への憧れ

AI（人工知能）とは「思考機械」のことである。「知能」を定義しようとするのが、世間一般ではコンピュータが人間のよう  
に考える日がきた（？）と信じられているのだ。だが、いったい機械が思考することなど、本来できるのだろうか。「思考とはそもそも、  
どういう行為なのか。ものを考えられるのは人間だけではないのか」——まず、そんな問いかけから始めるのも悪くない。AIロボット  
開発だのAIフィンテック（金融情報技術）だの、最新の技術動向がもたらす社会変化や経済効果をあたふたと追いかけているばかり  
では、かえって問題の根本的構図が見えなくなり、道に迷うことが多いからだ。

まず、思考が人間以外の存在には絶対に不可能だ、と断言することは難しい。たとえばペットの犬や猫にしても、それなりに何か考  
えているような気がしないだろうか。少なくとも、人間の気持ちを察して行動するなど、

喜怒哀楽の感情があるというのは飼い主のほとんど一致した意見である。思考とは論理的なものだけでなく、そこには身体活動に  
支えられた無意識の情動もはたらいている、というのは最近の脳科学研究が明らかにしたことだ。だが、動物に何らかの思考機能があるとしても、それは脳神経のある生きた身体を持っているからである。生きているわけでもなく、半導体回路の組み合わせでしかないコンピュータに、思考などできるはずはない、となってしまう。それなら、人工知能など、はじめから失敗を運命づけられた愚かな挑戦なのだろうか。

ここで、いったん、百年あまり前のヨーロッパの思想状況に立ち戻ってみなくてはならない。当時、論理主義という考え方が隆盛をきわめていた。アリストテレス以降の最大の論理学者と言われるゴットロープ・フレーゲというドイツ人が、『概念記法』という本を書いたのは19世紀末のことである。これは人間があつかう概念をふくむ様々な論理的文章を、記号を使って明快かつ厳密に

あらわすための方法をのべたものだ。この記法を使って、20世紀はじめ、英国の哲学者バートランド・ラッセルとアルフレッド・ノース・ホワイトヘッドが著した『プリンキピア・マテマティカ（数学原理）』こそは、論理主義哲学の金字塔である。人間のおこなう思考とは何かを正確に表現するようになる、というお手本が示されたわけだ（少なくとも、著者たちはそう考えた）。簡単に言えば、記号が何を意味しているかによらず、決まったルールにしたがって形式的に記号を操作して結論を導くのが「正しい思考」なのだ、という考え方である。まさにここで、記号の論理操作機械であるコンピュータのイメージが浮かび上がってこないだろうか。

コンピュータの発明は1940年代後半のことである。理論モデルをつくったのは英国の数学者アラン・チューリング、そのモデルをハードウェアに実装するアイデアをまとめたのはハンガリーから米国に亡命した数学者ジョン・フォン・ノイマンだが、兩人とも生粋の論理主義者に他ならなかった。だからコンピュータという存在は、当初からただの計算機械ではなく、そもそも高速の論理操作機械つまり「思考する機械」として誕生したのだ。人工知能（Artificial Intelligence）という用語は1956年のダートマス会議で最初に用いられたが、そこでは

『プリンキピア・マテマティカ』に書かれた諸定理の一部を自動的に導くAIが提示され、人々を驚かせたのである。

とはいえ、このAI第一次ブームは長く続かなかった。歯が立つのはせいぜいパズルやゲームなどで、実用的応用範囲が狭すぎたからである。その後1980年代になって、人間の知識を論理命題としてメモリに記憶し、専門家にかわって自動推論しようというエキスパート・システムが開発され、第二次ブームが起きた。日本の第五世代コンピュータ研究開発が巨額予算を使って行われたのもこの頃である。だが、予測に反して、医者や法律家が失業するわけでもなく、この第二次ブームも20世紀末には挫折してしまった。



---

## ビッグデータと統計処理

いったい、第二次AIブームはなぜ挫折したのか。第五世代コンピュータはなぜ失敗したのか。——種々の議論があるものの、半世紀以上コンピュータと付き合いしてきた私の目には、答えは明らかである。コンピュータは記号の論理操作機械だが、実生活における人間の思考の中で論理操作はごく一部でしかないからだ。医者や法律家の専門知識は確かに論理的に構成されているが、そこにはかな

らず柔軟な曖昧性が含まれている。さもなくば誤診も冤罪もない。英語の文章の意味には多様な解釈がありうるので、ただ機械的に翻訳しようとしても困難なのは当然である。

そして、2010年代後半になって、再び第三次AIブームが起きた。理由はどこにあるかと言えば、答えは簡単である。「間違ってもいいじゃないか」と考えを変えたためだ。コンピュータはかならず正解を導く思考機械だと見なしたために挫折したのだが、「まあ、だいたい合っていればいいじゃないか」と開き直ったのである。

第一次、第二次ブームのキーワードはそれぞれ「論理」と「知識」だったが、第三次ブームのキーワードは「統計」である（表を参照）。膨大なビッグデータを統計処理して、確率的に高い答えを出力するわけだ。これによって応用範囲は一挙に拡大されたが、一方、本来の思考機械の特長だった論理的厳密性は損なわれた。この点を忘れてはならない。たとえば、英文の機械翻訳では多数の用例をベースに頻度の高い翻訳文を出力するのだ

が、怪しげな誤訳も少なくないのである。

関連して指摘しておこう。第三次ブームの核心技術として有名なのは「深層学習（deep learning）」という方法である。囲碁の勝負でAIが世界的名人を破ったとして大ニュースになったが、そこでも深層学習技術が活用された。これは、パターン認識技術の一種であり、二つの特長をもっている。第一は、人間の脳神経系に類似した「ニューラルネット・モデル」を用いていること。第二は、従来のパターン認識技術と異なり、パターンの特徴を人間が与えるのではなく、コンピュータが自ら学習していくということ。第二の点は少々ふしぎな感じがするだろう。たとえば画像認識をするとき、画像の特徴を入力せずに、AIはいかにして画像パターンを把握するのだろうか。——答えは統計処理にある。たくさんの画像パターンを比較して、共通の特徴をもつパターン同士をグルーピングするのだ。こうして、グーグル社の深層学習ソフトは、ユーチューブから取り出した約1000万枚の画像から猫の顔を自動認識する

ことに成功した。

ただ、あえて付言すると、この猫認識においては、16台のプロセッサを連結した複合コンピュータ1000台を三日間つづけて

表 人工知能ブームの歴史

人工知能ブーム	キーワード	代表的応用	誤りの可能性
第一次 (1950~60年代)	論理	小 (パズル、ゲーム)	無し
第二次 (1980年代)	知識	中 (エキスパート・システム)	少ない
第三次 (2010年代~)	統計	大 (パターン認識、機械翻訳)	有り

作動させたとのこと。まさに途方もない計算量である。この例からもわかるように、第三次AIブームを起こした要因は、インターネットに蓄積された膨大なビッグデータ、そしてまた、それを処理する高性能コンピュータの出現にあると言ってよい。実は、深層学習のようなニューラルネットの理論モデルはかなり昔からあったのだが、データ不足と処理能力不足のために実用化されなかったのだ。その意味では、現在の第三次ブームはコンピュータ技術の漸進的進歩における単なる一段階であり、別に特別なものではないと考えられるのである。



## ◆ 神に近づく人間

冷静に見れば、現在は第三次ブームで盛り上がっているものの、憧れの理想的思考機械が出現したわけではない。ところが、一部のAI研究者からは、やたらに夢をふりまく声も聞こえてくる。深層学習によって、AIが人間のように自ら概念を把握できるようになったというのは、その典型である。これは明らかに過大評価というものだ。深層学習ソフトは画像や音声のパターンを統計的に分類するが、それが人間の社会的な概念と一致する保証はない。AIがたまたま外形的特徴か

ら猫の顔を分類できたとしても、もっと多様性の大きい犬の顔だったらどうか（ブルドッグやらシェパードやらチワワやら、顔はいろいろある）。AIは、生物的概念など全く理解していないのである。

深層学習の誇大宣伝は研究予算獲得のためかもしれないが、もっとも困るのは、声高に理想的な思考機械の出現を予言する「シンギュラリティ（技術的特異点）」仮説である。元祖は米国の発明家レイ・カーツワイルであり、その議論によれば、2045年頃に人間より賢いAIである超知性体が現れるという。このシンギュラリティ仮説の信奉者は欧米にはかなり多い。いったいなぜ、シンギュラリティ仮説のようなものが唱えられるのだろうか。シンギュラリティ仮説のベースには、前述の論理主義、つまり人間の思考とは基本的に論理的な推論であるという思い込みがある。確かに人間の思考には論理的部分もあるが、人間は動物の一種だし、大半の動物は情動をふくめた身体的直観によって生きている。身体的直観を無視して論理操作ばかり強調するのは見当違いではないだろうか。

しかし、人間は理性をもっており、ゆえに動物とは異なる特別な偉い存在だという考えは欧米では今でも根強いのである。さらにこういった理性重視の奥底には、古代から

のユダヤ＝キリスト＝神教にもとづく「ロゴス (logos) 信仰」がある。われわれ普通の日本人には分かりにくいのが、これはもっとも肝心な点なのだ。

ロゴスというギリシア語には二重の意味がある。第一はアリストテレスやフレーゲが研究した論理学の「論理」であり、記号論理自体は日本人にも理解できないわけではない。問題は第二の意味である。端的にいえば、ロゴスとは「神が受肉した存在」つまり「救世主キリスト」であり、したがって絶対的な「真理」をあらわす。ヨハネ福音書の冒頭部分「初めにロゴスがあった。ロゴスは神とともにあり、ロゴスは神であった」を想起して頂きたい。ユダヤ＝キリスト教では、全知全能の神が世界の万物を創造したことになる。人間も動物も機械もその点では同じ被造物であって、この被造物をつかさどる体系的な秩序原理がロゴスに他ならない。だから、もし機械がロゴスをうまく体現できれば、それが人知を超えた超知性体になるという理屈が出てきてしまう。

ここで大切なのは、全知全能の神と人間とのあいだの距離である。本来、両者を隔てる距離は無限に大きいはずだが、歴史的に段々近づいてきてしまったのだ。まずは四世紀のニケーア公会議で決められた、ローマカトリックの三位一体の教義がきっかけである。

そこでは神の絶対的真理を体現するイエスが、(聖霊とともに) 神と一体で同格の存在だとされる。人間イエスはユダヤ教の改革者だったのだが、イエスを介して、神と人間は一步近づいた。さらに宗教改革と、啓蒙思想による近代化が訪れる。これらをへて、絶対神への信仰よりも人間の理性が前面に出てくると、両者間の距離はいつそう縮まり、あたかも「人間が神に替わって世界を改善できる」という考え方が徐々に力を増していった。



#### AIは何をめざすべきか

AI技術に通暁したフランス哲学者であるジャン＝ガブリエル・ガナシアは、シンギュラリティ仮説を「現代のグノーシス神話」と批判している。グノーシス主義とは古代からあるキリスト教最大の異端だが、そこでは「人間と神が対等」なのだ。いかに神と人間が近づいたにせよ、キリスト教では神はなお人間よりはるか上位にある。しかし、キリスト教の素朴な信仰心が衰えて科学技術が称揚される現在、人間がいわば神と化して世界を改変していくべきだという思想は一種の常識として普及しつつある。そこに国際的情報産業の利害計算がからむので、話はいつそう厄介になるのだ。

私は信者ではないが、キリスト教自体は人々を救うすばらしい宗教に違いない。だが、その教義を巧妙に利用して人間の力量を過大に評価するなら、未来には悲劇が待っているのではないか。人間とは、生物進化のプロセスで偶然出現した、異常発達した大脳前頭葉をもつ奇妙な動物にすぎないのだ。AIに絶対真理につながる賢明さなど望むのは、暴論というものである。むしろ、欧米にも人知の限界性を理解している知識人は存在するが、欧米追随の日本人の中には、文化的背景を知らずにシンギュラリティ仮説に唱和し、大声をあげる情けない連中も少なくない。

端的に言って、いまのAIが約束するのはせいぜい、過去のデータを上手に統計処理して人間に提供してくれる自動機械の機能だけである。自動機械という意味では、エレベータとたいして違いはない。たまたま囲碁や将棋で名人に勝ったからといって、AIロボットが自律性を獲得し、自由意思をもつ賢明な主体になるわけではないのだ。SFめいた話はもう打ち止めにしてはどうか。

ただし、このことは決して、AI技術そのものの否定にはつながらない。過疎地の高齢者にとって安全な自動運転車は有効だろう。原発事故の後処理にもAIロボットの活躍が期待される。トンネルや橋など高度成長期

のインフラは老朽化しつつあるから、センサーで状況データを集め、早めに対策をとるためにもAI技術は不可欠である。

つまり、馬鹿げた幻想は捨て、あくまで現実的なコンピュータ処理の応用拡大として真剣にAI技術をとらえるべきなのだ。マスコミは幻想を振りまいているが、それはかえって第三次AIブームを短命に終わらせる結果になるだろう。

#### <参考文献>

- ・レイ・カーツワイル『ポスト・ヒューマン誕生』（井上健監訳、NHK出版、2007年）
- ・ジャン＝ガブリエル・ガナシア『そろそろ、人工知能の真実を話そう』（伊藤直子監訳、早川書房、2017年）
- ・西垣通『ビッグデータと人工知能』（中公新書、2016年）
- ・西垣通『AI原論』（講談社選書メチエ、2018年）

プロフィール……………  
にしがき・とおる 東京経済大学コミュニケーション学部教授。東京大学名誉教授。1948年東京生まれ。東京大学工学部計数工学科卒業。工学博士。日立製作所主任研究員、米国スタンフォード大学客員研究員、明治大学教授、東京大学社会科学研究所教授、東京大学大学院情報学環教授をへて、2013年より現職。文理にわたる基礎情報学の構築にとりくんでいる。著書に『デジタル・ナルシス』（岩波書店、1991年、サントリー学芸賞）、『基礎情報学』（NTT出版、2004年）、『ネット社会の「正義」とは何か』（角川選書、2014年）ほか多数。