

人工知能は人間の知能を超えるか？

—心理学のモデルとしての人工知能

立命館大学情報理工学部 教授

谷口忠大 (たにぐち ただひろ)

Profile—谷口忠大

2006年、京都大学大学院工学研究科博士課程修了。博士（工学）。2017年より現職。パナソニック客員総括主幹技師兼任。専門は創発システム論。著書は『記号創発ロボティクス』（講談社）など。「ビプリオバトル」発案者。



「人工知能は人間の知能を超えるか？」と、多くの人が第三次人工知能ブームのハイプの中で熱に靡かれながら問う。AIの囲碁プレイヤーは世界王者に勝利し、自動運転車の実用化は目前に迫り、機械翻訳の性能は上がった。過剰な流行に眉をひそめる専門家も多いが、技術的、学問的には地道な学術的努力に支えられた進歩があるし、機械学習理論は画像認識、音声認識、意思決定最適化、機械翻訳、自然言語処理等を支える基礎理論として成長を遂げてきた。

さて、人間と人工知能はどちらが賢いのか？ 賢くなるのか？ しかし、それより本質的な問いは「そもそも、僕たちは人間の賢さを理解しているのか？」である。

人工知能を作ろうとすると、私たちはイメージに描いた人間の「賢さ」を計算モデルとして具現化させることになる。タスクを限った時には様々な知能を作ることができる。例えば、タスクを「人間の顔認識」や「マリオブラザーズのプレイ」に限った時、僕たちは既に人間よりも上手くそのタスクを実現する人工知能を作ることができる。知能を機能として単離し理解した時、人間の知能を超えるような知能を作ることができるのだ。しかし、より総合的な知能を作るには、その「総合的な知能」が何物なのか、僕たちはま

ずそれを把握しないといけない。

よく出くわす「あるある話」だが、人工知能の研究に興味のある学生にやりたいことを聞くと「人間みたいな知能を作りたい」「人間と対話できるシステムを作りたい」「意識と感情を持つロボットを作りたい」という。壮大な野心だ。しかし「では、何を示せば、それらが作れたことになるのか？」と問い返すと、多くの人は答えに詰まる。工学的な願望は、壮大な哲学的、また、心理学的問いへと姿を変える。

どんな人間より早く走れる自動機械は二世紀前にできている。どんな人間より高速に計算を解ける自動機械は皆が持っている。どんな人間より強い囲碁プレイヤーもできた。しかし、これらの「証明」を持ってしても、「人工知能が人間の知能を超えた」ことの証明にはならない。なぜならば「人工知能は人間の知能を超えるか？」という問い自身が適切に定義されていないからだ。そしてこの問いはこれは、僕たちに人間の知能理解に関する挑戦を突きつける。つまり「人間の知能とは何なのか？」—ここから皆さんを人間の知能理解の無限回廊、いや、螺旋階段へと案内したい。

研究者、探求者、世界の理解者としての人間が物事を理解するにはいくつかのステップがある。現

代の科学的研究の多くは方法論として実証主義的な考え方を基礎としている。何らかの「仮説」を立て、適切な実験系を構築した上で、実験データと統計的仮説検定により、その真偽を確かめる。仮説は反証可能性を担保しなければならない。その中で検証に耐えた仮説だけが記述可能な知識、つまり、形式知として蓄積されていく。

しかし、形式知の集合体が私たちの知、または、現象理解の全てかということ、それは違う。形式知をいくらデータベースとして保管しても理解にはいたらない。理解とは暗黙知である。また、実証研究においては、仮説自体を形成するプロセスは一般的な手続きの外側にある。実証という帰納的思考だけでなく、演繹的思考やアブダクション的思考が不可欠だ。

人間の知能理解に、非常に重要な存在でありながら、このような手続きの外側にあるのが「モデル」だ。モデルとは人間が対象系を把握するための構造的な表現である。物理学では多くの優れた数理モデルを提案し、実証研究により淘汰し洗練することで、進歩を遂げてきた。豊かで優れたモデルは現象理解を深める。物理学のみならず多くの学問においてモデルの役割は決定的に重要だ。

モデルとは一方で色眼鏡である。人は物事を理解する時にモデ

ルを通して理解する。ネガティブな意味においては知覚を捻じ曲げることもあるし、ポジティブな意味では世界の理解を深める。

モデルという言葉とアナロジーという言葉は親戚のような存在である。モデルは自らが知る現象が構造的に抽象化されることで、その個人により理解されている場合が多い。これをアナロジーで研究対象に写像することで、僕たちは物事を把握する。

僕は「人間機械論」という言葉が好きだ。その考え方が好きなのではない。その考え方の背後にある示唆が好きなのだ。ニュートン力学やそれに基づく制御工学を基礎とした世界観を持ち、外界に多くの自動機械を観察し続ける現代社会人が、人間について理解しようとする時に、「人間も結局は機械のようなものではないか?」「人間はロボットと変わらないのではないか?」といった素朴な人間機械論が立ち現れる。これは、自らが「見てきたもの」「仕組みを理解したもの」をアナロジーの元にして人間という対象を把握しよう、理解しようとする人間の自然な認識活動である。つまり、機械やロボットをモデルとして人間を理解しようとしているのである。

「人間はロボットではない、何か違う『人間特有のもの』を持っている」という人もまた、ロボットというモデルと、そのロボットというモデルではモデル化していないモデル化誤差としての「人間特有のもの」という把握をしている時点で、しばしばロボット・機械というモデルに囚われている。

結局のところ私たちは物事を理解する時に、自分たちが目にしていて、自分たちが作り出したもの、自分たちが作り出した数理モデルをアナロジーの元、モデル

の元に置くしか無い。人間理解を豊かにしたければ、人間理解の時に人間に当てはめるモデル自体が豊かにならなければならない。人間の知能理解のための人間のモデルとは、人間の知能を人工的に外在化させたものとなる。それが「人工知能」なのだ。

僕自身が推進している研究分野は人間の認知発達や言語獲得を機械学習やロボティクス技術を用いて再現し検討する学問分野であり、「記号創発ロボティクス」と呼ばれる。「記号」とは言語を含む表象の総称であるが、人間の知能を考える際に感覚運動系に基づく実世界認知と、他者との社会的なコミュニケーション、また、理性的な思考を媒介する存在として極めて重要になる。環境に適応する中で、環境を理解し、言語を獲得し、他者とコミュニケーションを図る。そのような「やわらかい」知能のモデルを作ることが記号創発ロボティクスのチャレンジである。

「人間はロボットではない、何か違う『人間特有のもの』を持っている」という主張に関する、ナイーブなアプローチは、その差分に焦点を当てて「人間特有のもの」を探究することである。しかし、これは、あまり筋が良くないように思われる。

二つ理由がある。第一に「人間特有のもの」の理解にもモデルが必要であり、そのモデルは、ロボット側のモデルと結果的には融合し、人間の構造的、いわば「ロボットの」な理解を導いてしまう。これを避けようとする、結局は曖昧な人間理解にとどまってしまう。第二に、ロボットが人間の努力により前進する科学技術の産物であることを忘れていて。科学技術は進歩する。ロボットとい

う存在の変化によって、人間とロボットの境界は変化していく。21世紀前半での「人間特有のもの」は、23世紀の「人間特有のもの」とは全く異なるであろう。19世紀において「人間特有のもの」が何だったか想像することは良いトレーニングになるだろう。

現時点では異端かもしれないが、王道たるアプローチは「ロボット」の可能性を広げ、その意味を拡張していくことであろう。そして「人間特有のもの」の領域をとにかく狭めていくのだ。ロボットが僕たちにとって自然なまでに柔軟でふくよかな存在になれば、人間理解のモデルとして、それは、より豊かなものになるだろう。

「人工知能は人間の知能を超えるか?」——この問いに答えるには「人間の知能」をより良く把握、理解しなければならない。つまり、心理学が発展しなければならない。そうしなければ、この問いがよく定義されない。

結局、人間の知能理解を深め、人間の知能が人工知能に負けないことを示すためにも、人工知能を人間の知能に接近させる最大限の努力をしなければならない。人工知能の研究により人間の知能のモデルが発展し、そのモデルにより人間の知能理解が深まれば、人工知能の研究もまた発展する。

自己言及的に思われるかもしれないが、これは同じ場所をぐるぐる回る、出口の見えない無限回廊ではない。繰り返しによって行ったり来たりしながら人間の知能理解の高みへと駆け上がっていく螺旋階段なのだ。「人工知能は人間の知能を超えるか?」を反証可能な問題に変えていくことこそ、これからの心理学のチャレンジなのだ考える。