

AIに心はあるのか



デジタルエコノミー推進委員長
日本電信電話相談役
しのはら ひろみち
篠原 弘道

最近、数十年ぶりに再読している本がある。ロジャー・ペンローズの『皇帝の新しい心』である。著者は、数種のタイトルで非周期的に平面を埋め尽くす「ペンローズ・タイル」やブラックホールの特異点定理など、数多くの業績を挙げた数理論理学者である。

この本にもう一度挑戦しようと思ったきっかけは、最近のAIをめぐる議論にある。特に、生成AIの急速な発達に伴い、AIは人を超えるという論調なども含め、「AIと人」という対比を耳にするたびに、「AIは『人の心や意識』に相当するものを持ち得るか」という点と、「生成AIがどのように言語獲得を行っているのか」という点についての疑問が頭に浮かんでくる。

この本は、副題に「コンピュータ・心・物理法則」とあるように、「AIは心を持っていると言えるのか。心とは何か。意識とは何か。」というテーマを、30年以上前に論じた一般向けの書籍である。著者は、「コンピュータは心を持ち得るか」という問いに対して、心や意識を物理学で説明しようとする展開している。そのために、チューリングマシン、ゲーデルの不完全性定理、フラクタル、複雑性、量子力学、時間の矢、脳科学など、様々な切り口から人の心の動きにアプローチしている。本書では、量子重力理論が人の心を説明するための鍵ではないか、という考えを提起してい

るが、残念ながら人間は単に「肉体を持ったコンピュータ」なのか否かの答えは示されていない。しかし、様々な分野の知識を総動員して問題にアプローチする著者の姿に対して畏怖を覚えたことが、鮮明な記憶として残っている。内容をきちんと理解したとは、決して言えないが、様々なテーマに興味を抱かせた本であり、私にとっては重要な意味を持った本である。

最近、ある著名なAI研究者とLLM(大規模言語モデル)について会話した際、AIの心・意識の問題、言語獲得メカニズムについて、私の疑問をぶつけてみた。彼は、「LLMに関する研究の進展が、人間の言語獲得メカニズムを明らかにする可能性があるかもしれない」と語っていた。果たして、生成AIの今後の発展が、人間の言語獲得能力について何らかの知見をもたらすことはあるのだろうか。AIが人間と同様の心や意識を持つことは、今後も不可能だと私は思っているが、生成AIの発展と認知科学の融合によって、言語獲得の謎が明らかになる可能性がゼロではないと考えるとワクワクしてくる。



『皇帝の新しい心
コンピュータ・心・物理法則』
著者：ロジャー・ペンローズ
みすず書房