

講演要旨

AI であるとはどのようなことか

帝京科学大学学長

沖永宜司

この発表では今日、目まぐるしい勢いで高度化する AI の姿が、すでに人間の推論能力や対話能力と機能的には変わらず、部分的には人間をはるかに凌駕する段階まで達していることに鑑み、AI に代替できない人間の性質は残り続けるのか否かを考察する。そして少なくとも機能上は AI に代替不能な人間的な働きは究極的には存在しない可能性があり、しかしそれは反対に、私たち人間の主体性や人格なども実は最終的には非実体的であったことを明らかにして行く過程でもあることを確認する。確かにその過程は人間存在を否定し無価値化する危険性を含んでいる。しかしそれは反対に、AI によって将来実現されるかもしれない私たちの意識をも含んだ非実体的な情報ネットワークそのものが、新たな普遍的な生命の姿にもなり得る可能性をも示唆している。

操作されるコンピュータとそれをプログラムする人間主体との根本的な区別という「伝統的」な構図は、コンピュータ科学者の間では崩れつつある。AI の CPU 全体に人間的な主体の役割をも見出されているからである。それは彼等が情報ネットワークから主体や意味が「創発」することをこれまで素朴に受け入れてきたからであり、そこで一人称と三人称などの存在論的な区別の消失も起こっている。「創発」とはこの世界の中にある種の「不合理」を認めることであるが、この「不合理」は実際の主体的役割の実現というプラグマティックな理由から是認されてきた。そこでは AI の予測不能な行動に AI の自発性と意識が見出され、客体でしかなかったコンピュータが「主体」化する瞬間も描き出される。

この汎主体的構図は一見非人間的であるかもしれない。しかしこれは、同じ不可分の経験の一部が、あるときは主体となり別のときには客体となる人間の意識の性質と類似した構図にほかならない。主体とは特定の位置にある実体ではなく、そのつど役立つ機能的概念だからである。人間の片方の脳半球も主体となり、客体にもなる。1つの脳細胞も、または巨大なネットワークに接続された人間の脳も同様に主体となり客体となる。このように人間が CPU の主客両者になり得ることは、超越論的主体という世界の例外の否定であり、徹頭徹尾世界内存在であることの再確認とも解釈できる。

こうなると情報システムとして AI のネットワークは人間の知能に機能上ほぼ一致する。しかし AI には人間の意識と同じ内面性が伴うのか、という疑問も呈示されるだろう。これはコンピュータに意識をアップロードすることへの疑念をもたらしものである。そしてこの疑念は、意識が機能と物質のどちらに基づくか、という問題に関係している。するとマイクロチップと電気信号を基盤とする現状の AI に対して、培養された脳細胞を基盤とする脳オルガノイドが AI に応用された場合、機能と物質の両方が人間の意識にとって代わる可能性が生じる。このときその AI の内面性、その AI であることがより現実味を帯び、AI と人間意識との内的連続性が、より深刻な問題としても提起され得ることになる。