SS10　AI と人間の知能の二つの根本的差異

Two fundamental differences between AI and Human Intelligence

大澤真幸

社会学者

Masachi Ohsawa

Sociologist

=抄録=

本講演は、人工知能 AI と人間の知能との間にある〈根本的差異〉について論ずる。この〈差異〉は、医療に AI を（どのように）使用すべきか、という主題とは直接には関係がない。だが、それは、AI を活用する者が理解しておくべきことである。AI と人間の知能の間には、架橋することができない〈差異〉が二つある。ここで「架橋不可能」とは、人間には備わっているが、AI が、現在前提になっているアイデアのもとでいかに発展し性能を上げても、決して獲得できない能力という趣旨である。

第一の〈差異〉は「フレーム問題」に関係している。フレーム問題とは、行動を遂行する際に、いかにしてレリバントな事項をイレリバントな事項から（効率的に）区別し、選択するか、という問題である。人間がフレーム問題を克服できるのは、イレリバントな事項を〈無視する〉ことができるからだ。（知っていることを）無視することができる AI をどのように作ればよいのか、見当もつかない。なぜなら〈無視〉とは何もしないことであって、積極的に実行してしまえばもはや〈無視〉ではないからだ。ChatGPT のような大規模言語モデルは汎用性が高いので、フレーム問題を克服しているかのような印象を与えるが、それは、膨大なデータセットの事前学習と計算速度の高さによって、〈無視〉と機能的に等価なことを積極的に実行しているからだ。フレーム問題を解決しているわけではない。

第二の〈差異〉は、「記号接地問題」に関係している。人間は、記号（≒言語）を外部の実在や世界に結びつけている。人間にとって「記号の理解」は、〈記号接地〉されているということを不可欠の条件として含んでいる。しかし AI は記号を操作する機械なので、その記号は外部の世界や実在に接地させてはいない。

人間にとって、〈無視〉も〈記号接地〉も難しいことではない。何かをやる前に「すでにできてしまっていること」である。しかし AI には、それらは原理的になしえない。どうして、人間にはそれらができるのか。その鍵は、人間の心の独特の〈社会性〉にある。社会学者が AI に関心をもつ所以はこの点にある。本講演では、これらの能力が、人間固有の〈社会性〉とどう関係しているのか、私の仮説を提示する。