

人工知能に心はあるのか？

小笠原 義仁 (Yoshihito Ogasawara)

早稲田大学 総合研究機構 複雑系高等学術院研究所
芝浦工業大学
三芳合金工業株式会社

近代科学は数学を言語としていることにより特徴付けられるが、それは科学が応用数学であることを意味する。ここで、応用数学（科学）には修辞法が使われていることに注意されたい。実際、応用数学によって新たな類似が創造されて、ある細部は隠されて、ある細部は強調されることにより、世界観が構築される。それは、洞察（インサイト）を与えるものであり、意味の創造である。記号とは本質的に隠喩的であり、「記号」と「意味」との完全な一致など有り得ず、両者の間には絶えず葛藤がある。例えば、言うまでもなく記号「 x 」は記号「 x 」自身に過ぎないのであり、物体の位置として解釈されていたとしても[1]、物体の位置そのものでないことは明らかである。その時（単なる）記号「 x 」は物体の位置に隠喩されている。すなわち科学とは、現象と数学とを隠喩によって結びつける営為であり、科学において「数学を言語として用いる」という営為は文学であり詩でさえもある。そこで要求されているものは、記号「 x 」を物体の位置「として見る」という行為であり、読者の想像力であり創造力である（それは芸術である）。一方で科学の言語である数学は、厳密な論理性も要求されている。従って科学は、数学を言語とすることにより、芸術性と論理性との調和が図られることになり、科学の魅力や科学の審美性は、まさにここに起因する（実は数学自身が持つ審美性も、この芸術性と論理性との調和に起因する）。

一方で科学は、客観主義の名のもとに「私」の存在を隠蔽してきた。言うまでもなく、世界は「私」の意識から生み出されるものであり、従って世界は意識に依存するが、科学はこの事実を無視し続けることにより、世界の存在を自明なもの、アприオリなものとしてきた。さらには、その世界観からあらためて意識の問題を問うという因果を逆転した議論をすることにより、疑似問題を捏造することにすら成功している。

本研究では、数学を言語とするという科学の特徴に倣いながら、意識から生み出される世界を描くことを目指している。それは、客観主義の名のもとに、自らを観察者の立場に逃げ込ませるのではなく、認識をしている自らの認識の仕方に眼差しを向けて、そこから目を離さないことによって始まる冒険への第一歩を踏み出すことである。またそれは、科学において失われた「私」を奪還するための闘いであり、客観性の名のもとに形骸化された世界におけるレーベンの叫びである（すなわち実存の闘争である）。一方でそれは、タブラ・ラサにおいて真っ白なキャンバスに、いずれ美しい絵画になることを夢見て、試行錯誤をしながら下書きを繰り返す行為であり、豊かな果実を求めて未知の島を探す探検である（それは新しい科学を探求する冒険である）。

特に本研究では、言語である数学としてトポロジーを用いている。トポロジーは形態の概念を記述する言語であるが、本研究では上記の目的を果たすために概念の形態（すなわち概

念の概念)を記述する言語としてそれを利用する[2,3]。また世界のモデルとしては、Primitive Chaos [4]と呼ばれる概念を利用する。ここでは、世界は決定論と非決定論とが不可分にもつれあっているものとして記述される。

なお、科学は決定論的世界観を前提とすることが多いが、それはそれが予測可能性という実用性を生み出すからである。しかし、世界が本当に決定論的であるとするならば、それは我々を絶望に突き落とすものである(人工知能が決定論的プログラムのみで限界があったことは示唆的である)。実際、決定論的世界はラプラスの悪魔と呼ばれる世界観を導く。すなわち、本来世界は決定論的であり、全ての法則と、ある瞬間における全ての状態を知っている知性(悪魔)が存在するならば、不確実なことは何もなく、未来についても完全に予想することができるという世界観である。それによると、例えば我々が自由意志により選んだ人生と考えるものも幻想に過ぎず、そもそも自由意志なるものは我々の認識不足に起因する幻想に過ぎないということになる。この絶望的な世界観において、しかし上記のように全知と呼べるような知性は、「悪魔」と呼ばずにむしろ「神」と呼ぶべきではないのだろうか。このことに気が付くならば、例えばキェルケゴールによる次の言葉に我々は突き刺される(「死に至る病」斎藤信治訳、岩波文庫、2023)。

『雀が生きていられるのはなるほどもっともである。雀は自分が現に神の前にあるということを知らないのだから。だがひとたび自分が現に神の前にあるということを知った人間がその同じ瞬間に発狂もせず破滅もしないということがありえようか!』

決定論的世界観は絶望に向けての疾走である。実用性を得るために、戦略的に決定論的世界観を採用するのではなく、「本気で」決定論的世界観を信じた時に「私」の存在は崩壊する。多くの科学者がそのような崩壊をもたらさないのは、科学がもつ「私」の秘匿性に守られているからである。実際科学は、観察者と呼ばれるシェルターに「私」の存在を隠匿することにより、決定論的世界観と呼ばれる核爆弾から身を守ることに成功している。しかし、この科学がもつ欺瞞に気付いたならば、シェルターなど本当は無いことに気が付かされることとなり、たちまちその身は焼きつかされる。

Primitive Chaos においては、決定論的世界観は前提とされておらず、ただ認識可能な現象があるのみである。しかしそれは時に、予測可能性を生み出すものでもあり、決定論的世界観と非決定論的世界観がエンタングルされたものである。このような Primitive Chaos と呼ばれる概念を、トポロジーと呼ばれる言語で語らせることにより、我々は「私」の前に立ち現れてくる世界の描像を見る。しかもその時、「私」とは異なる「何か」からも世界が立ち現れてくる描像を見る。それは人工知能が持つ心なのであろうか。

[参考文献]

- [1] Y. Ogasawara et.al., Journal of the Physical Society of Japan, 72(2003)1871; 73(2004)1703; 74(2005)2439; 75(2006)64003; 78(2009)57001; 80(2011)113002; 81(2012)77001.
- [2] 小笠原義仁, ころの科学とエピステモロジー, 3(2021)6; 3(2021)48; 4(2022)46; 4(2022)71; 5(2023)118; 6(2024) in print.
- [3] 小笠原義仁『ものの見方としての位相空間論入門』培風館, 2011.
- [4] Y. Ogasawara et.al., Journal of the Physical Society of Japan, 79(2010)15002; 80(2011)67002; 81(2012)103001; 83(2014)14001; 84(2015)64007.