

強い人工知能へのキリスト者である工学者からの一考察

堀江幸生 (Sachio Horie)

名古屋大学 情報学研究科 博士後期課程

概要 (Abstract)

The view that there is an inherent theological conflict between strong artificial intelligence, on the one hand, and biblical teaching regarding the origin of the soul, human worth, and humanity being created in the image of God, on the other hand, is examined and shown to be ill-founded. Christian theology, therefore, has no stake in the claim that the possibility of technological accomplishments in this area is inherently limited. Consideration is also given to how a biblical understanding of human personhood can inform work in artificial intelligence, and suggested that the ethical consequences of strong artificial intelligence need to be carefully considered.

Keywords: God, Man, Creation, Artificial Intelligence, Machine, Christian, Limitation

1. はじめに

人工知能は、我々の生活基盤のひとつとして機能している。我々は、我々の個人的、社会的行動と相互作用して強化される人工知能に責任を持たなければならない。子育てには子どもの行動に対する道徳的責任が含まれるのと同様に、我々は人工知能に対しても同様の責任を負う必要がある。人工知能に学習させる（教える）内容はしばしば、社会集団間の非道徳的な事項である。最近のある事例では、悪意のある人々が人種差別的行為を広めるために人工知能を調整した。

計算機科学に携わるキリスト者は、人工知能の未来を築かなければならない。多くのキリスト者である計算機科学者にとって、コンピュータ産業におけるキリスト教の役割は、倫理と道徳についての初等教育を提供する機会となると考えている。

人工知能について明確に考えるためには、我々が知能の意味を明確に考える必要がある。人工知能の歴史の中で計算機科学者は常に、人間のように考え、行動し、人間よりも合理的に考え、合理的に行動する機械の知能とは何かを問い続けている。

強い人工知能の実現可能性をキリスト者は考察してきた。強い人工知能の実現は、神によって吹き込まれた魂や人の創造など、人に関わるキリスト教の教義とは相容れない。福音主義的な神学者は、「強い人工知能の出現によって、我々の神学は完全に崩壊する」と述べている。

聖書の教えと強い人工知能と間には、決定的な対立が存在するであろうか。または、キリスト教神学には、そのような対立にアプローチする手法はあるのだろうか。

本稿では、キリスト者の立場から、強い人工知能を作ることが可能であるかどうか、あるいは、キリスト者が、強い人工知能の実現可能性を、特に以下の部分に焦点をあてて、どのように追求するかについて考察する。

1) 人の魂の起源をどのように考えるか。

2) 人は神によって創造されているのか。

3) 聖書の教義は、強い人工知能の研究に影響を与えるか。

子を持つ親であるならば、「なぜお母さん（お父さん）は、私を愛していますか？」、「あなたは私の兄弟（姉妹）を、私より愛していますか？」このような問を受けることがある。

詩編の編纂者は、似たような問を神にした。「そのあなたが御心に留めてくださるとは人間は何ものなのでしょう。人の子は何ものなのでしょう。あなたが顧みてくださるとは。」（詩篇 8:4）、詩編ではこの質問の答えはないが、新約聖書でイエスは究極の隣人愛で応えた。我々は、神との関係をすでに持っている。

2. 人工知能と意識

この章では、工学的見地ではなく、キリスト者の見地より、人工知能の意識を考える。

キリスト者にとって、自然知能の頂点にあるものは神である。創世記 1: 27-28 では、

「神は御自分にかたどって人を創造された。神にかたどって創造された。男と女に創造された。神は彼らを祝福して言われた。「産めよ、増えよ、地に満ちて地を従わせよ。海の魚、空の鳥、地の上を這う生き物をすべて支配せよ。」

現在、実用的な人工知能とよばれるものは、専門領域で機能する弱い人工知能の集合である。例えば、同じ深層学習でも画像認識と金融工学では解析するデータによって出力が異なる。

弱い人工知能においても、ある作業を機械に委ねることによる重大な倫理問題を引き起こす可能性はもちろんあるが、本稿では、強い人工知能が人の本質に関連する問題を提起することを鑑みて、強い人工知能（以下、人工知能）に焦点をあてる。

キリスト教では、知能の存在を人ではないものにも認めている（例えば、神と天使）。ここで、人とは何か（本質）を定義する。Rudder Baker の定義を借りれば、

- ・人を人とするのは、一人称の視点をもつ能力である。

彼女は、これを「他者とは異なる主体として、世界に直面する個人として自分自身を考える視点」と定義している。「全ての知的な存在は意識を有するが、すべての知的な存在が、一人称の概念を有するわけではない」と主張する。では意識とは何であろうか。意識とは何かを特定できれば、それを機械に与えることが可能であろうか。確かに我々は意識を身に付けていると答えることは可能である。意識というシステムは、人が自然から阻害された瞬間に産まれたとするならば、このとき、人は脳の中で、「世界（自然界）」というモデルを得たのである。そして、人は自分が自然の外にいることに気づき、不安であるがゆえに信仰をもった。これは、創世記にアダム の墮罪として象徴的に語られている。

意識を情報処理システムに置き換えることは、情報処理が、意識の持つひとつの側面でしかないという考えを越えて、意識というシステム全体が情報処理となる。意識は、感情や欲求、欲望、意志など様々な部分を持つが、これらを情報処理における副作用であるとは言い切れない。これらの部分が分割不可能なままに統合されて初めて意識と呼べるものになりえる。

ここでは、図 1. を提示する。各世界を核世界、自然界、機械世界、仮想世界と呼ぶ。重要なことは、核世界は自然界を生み出し、人が毎日感じている自然界の形と秩序を生み出す源泉であり、人を導く神の超自然の世界でもある。人工知能は、機械世界と仮想世界において人と相互作用的に存在する。

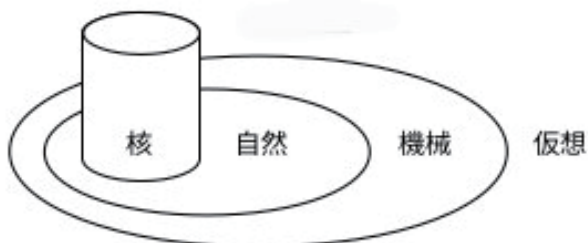


図 1. 世界モデル

3. 人工知能と人、そして神

キリスト教の教義と人工知能が対立することは、チューリングが現れてから顕著になった。

我々が、現在「人工知能」と呼んでいるものを唱える最も初期の論文は、“Computing Machinery and Intelligence”である。

キリスト教では、意識とは魂の機能だと考える。神はすべての男女に永遠の魂を与えたが他の動物や機械には永遠の魂を与えなかった。したがって、動物や機械は考えることができないことになる。この見解は、魂のそのものではなく、むしろ魂の起源に関係している。それは、神の創造的な行為が魂を通して物質的であるものとならないものとを区別するということである。技術は前者の

領域に属しているが、神のみが創造することができる機能ではなく、我々が可能な機能である。

このような理解からは、聖書の教えと技術の間には対立しか生まれないが、この理解は本当に聖書の教えであるのか。

人の創造は、創世記 2:7 に記されている。

“主なる神は、土（アダマ）の塵で人（アダム）を形づくり、その鼻に命の息を吹き入れられた。人はこうして生きる者となった。”

多くのキリスト者はこの節を、神が2つの別々の創造的な行為をしていると理解している。すなわち、神は人間の身体を最初に形成し、次に神は人に魂を吹き込んだ。この魂は、キリスト教では人の不可欠な構成要素だと理解されている。人の身体を創造することは、おそらくは神の即時行為、あるいは自然選択による進化を介する工程と見做されるが、魂を吹き込むことは物質的な領域外にある神聖な行為を伴うものと見做される。それは最終的に、神が“人の身体”を創ったということではない、神は、“人を創った”のである。

ここで我々は、創世記 2:7 における「人」を「人の身体」、 「息」を「魂」、 「なった」を「受けた」と考える。

人を特別なものにするのは、我々がなぜ神から創られたのかと言う目的に基づく“神との関係”である。キリスト者は、“脳”と“魂”という物質的な世界と非物質的な世界を、神が認めながら創造されたものとする。

我々は、人工知能の創造者として、それを神からの創造的な贈り物のひとつだと捉えたい。Phillip (2006) によれば、聖書は、我々と神との関係を定義している。すなわち、我々と神との関係とは我々が神の前で、どのように知的に振る舞うべきかではない、神から与えられた人生をどれくらい神に依って生きるべきかである。キリスト者にとって神との関係は、人生に意味と目的を与える。それは神（創造主として）と人（創造物）との間の契約である。では、コミュニケーション能力を持つ人工知能は、神との関係を確立することができるのか。

機械は人の製作した製品であり、人に替わるものではない。

McKerrow (2006) は、意識を持つロボットは、知覚で感知することを利用して、その周囲を知ることができ、障害物を認識し回避することができる」と指摘したが、ロボットが意思決定に道徳的責任を負う意識の欠如、すなわち、自分の行動に責任が伴うことがわからないのであれば、それが人と機械を区別するひとつの要因となり得る。ロボットは、人生の貴重な有り様、美しさ、完全性、正義について学ばなければならない。それこそが、神との関係を確立することであり、意識の源である。

人が何かをふと気づく時、本稿では、人のもつフレームがある瞬間に広がる（狭める）事象であると考えますが、それは脳の機能として十分に説明されてはいないように、我々は、脳がその部分的な性質の単純な総和にとどまらない性質を意識と言うシステムとして備え持っていると考え。そのシステムは、アルツハイマー病などの脳損傷および疾患により不完全になり、また、睡眠中または麻酔下で一時的に中断する。意識は、別々の起源をもつ脳と

魂からなるのではなく、物質的な脳に強く依存することを認めざるをえない。それでは、脳と魂、意識との関係性とは何であるか。

まず、アダムの子孫の人間の罪の普遍性は疑う余地がないとするならば、人の魂は、アダムより引き継がれたものとなる（キリスト教伝統主義）。キリスト教伝統主義と人工知能は、アダム（人）の魂の起源をどのように説明しているかについての見解は大きく異なっている。創世記の2:7が、物質的な行為を通しての非物質的な魂の伝播を説明している。人が意識をもつ過程は、十分に機能する脳とそこに吹き込まれる魂からなるはずである。しかしながら、人の意識が複雑な技術的な人工物の操作と同様の方法で現れることを否定できる理由がないことは自明であろう。ただし、意識の創造が、技術的に可能であると言うことは、実際に起こる可能性があると言うことを証明するものではない。例えば、人が何かをする時に採るような行動の目的を、人は話することができる。しかしながら、行動の結果、人の意識では問えられない副作用が生じることがある。それは歩いた時に、絨毯の上にいた虫を殺してしまったことかもしれない。このような副作用は人の行動とは切り離されているが、これが重要なポイントとなる。人が自分が行動することで引き起こした現象を知り得ないならば、機械がそうでないと何故言うことができるのか。機械は目的を持って行動しているかは見かけからはわからないが、少なくとも人は行動する目的を持っていると説明できる。人は意識して行動できる。これは魂に依るものであり、人の魂は、神との関係において必要不可欠なものである。

4. 人工知能、人の価値と神の想い

歴史的に、技術的な革命が起こる時、人はそれを、人の価値への挑戦と考える。計算機がある領域で、人よりも明らかに優れている場合、計算機は「単なる機械」であると決めつける。例えば、Garri Kasparov は、ディープ・ブルーに負けた後、「彼（ディープ・ブルー）は、勝利にもかかわらず、勝利を楽しむわけではないし、満足感を得ることはできなかっただろう」と語った。

人の価値が、人の知能と結びついているならば、強い人工知能の可能性は、人の最も重要な概念に深刻な脅威をもたらす。しかしながら、人は無条件の愛をもっている。人を特別なものにするのは、我々を創った神の目的に基づいた神との関係に他ならない。キリスト者は、人の価値を全く異なる視点から見ることが可能である。人の価値は、人の目的「神との関係であり、それは神が我々にさせたいこと」と関係していると我々は考える。すなわち、人とは何であるか、人の生物学的な性質は、人を他の動物と切り離して特別なものにしていないがゆえに、むしろ我々が人であるためには、何か特別なことができるようにしなければならない。

キリスト者は戸惑うであろう。いつか人を超えるような人工物が存在するのならば、それはキリスト教にとって何を意味するのか、人は神の想いで創造された価値あるものなのか、墮落した人が神の想いと同様な機能を有する機械を創造できるのか。神の想いの中に人工知能は用意されているか。この問いに対する答えの1つめは、人工知能と神の想いが、全く関係ないとするならば、キリスト者にとって人工知能は否定されなければならない。2つめは、人工知能と神の想いが、意識における人の理性と呼ばれる性質に関係するのならば、人工知能はキリスト者にとって肯定されるが、

それに依り頼むことは偶像崇拜となる。最後に、意識と神の想いは2つめと同様に関係しており、人工知能が人にとって、様々な問題解決の道具として存在するのならば、人工知能と神の想いの間には聖書的な矛盾はない。人の技術が人に等しい（あるいは場合によってはそれを上回る）知能をもつ機械を創造する時は、神が機械のために贖いを提供することを選択したのであり、結果として、人間が経験できる以上の関係を神ともつことができると仮定する。

人工知能が、地球上の支配権を行使する上で、人とパートナーであり、神について何かを明らかにすることができるならば、これが人の価値を無効にすることにはならないと考える。

最終的に、神に対する人の価値は、永遠に失われることはない。もちろん、人の創造主である神を参照せずに人としての価値は理解できない。強い人工知能の存在は、神の想いで創造された人の価値とは別に、人が神を信じないような部分で人の価値を見出そうとする試みを損なうだろう。神に祈る必要性を減少させる技術と時代は、神の想いを前にして人がどれくらい価値あるものなのかを隠してしまう。

5. 終わりに

聖書は、人工知能の研究に影響を与えるだろう。聖書の教えと強い人工知能との間に、神学的な対立があると主張する立場に、本稿では、対立軸を神の視点に置くことで再考した。強い人工知能の出現を聖書は妨げるものではない（強い人工知能の倫理的な影響は依然として慎重に考慮する必要がある）。それでは、人工知能について、聖書的な解釈から何を期待すべきか。

聖書は人を神の創造物の一部として描写しており、それ以上でも以下でもない。聖書は、人を「塵」とする。神は、人に抽象的な知識を与えるのではなく魂（意識→知能）を与え、そして、神は、物理的な脳を作り出した。これは、明らかに進化論的なプロセスである。強い人工知能は魂はもたないが、いずれは、脳をもつようになる。

人の起源をたどれば、創世記3章は、死が人による罪の結果であることを明らかにしている。しかしながら、その後、聖書は人への神の計画を記している。それは、神との永遠の関係、救済である。

人工知能の研究は、聖書で明らかにされた希望のひとつである。例えば、ある程度強い人工知能は、人類を長く悩ませている問題の解決を可能にする。もちろん聖書は、偶像礼拝を許していない。偶像崇拜に対しては、基本的に自分の必要を満たすために創造主以外の誰か、人物を探すことでである。

科学全体を通してそうであるように、人工知能の開発は、上記のように間違っただけ動機づけられるかもしれないが、本稿では、創世記1章28節、

神は彼らを祝福して言われた。「産めよ、増えよ、地に満ちて地を従わせよ。海の魚、空の鳥、地の上を這う生き物をすべて支配せよ。」

キリスト者は、人工知能をより深く理解することによって、神の創造物のなかでも最大の驚異を任務として管理すればよいとする。

最終的には、聖書の見解と人工知能との間には本質的な神学的葛藤はなく、人の価値を神の想いとして捉えればよい。これまで達成されてきたことを現実的に評価することは、近い将来達成される壮大な予測を可能にする。我々は、この領域に大きな注意と敬意を払って取り組む必要がある。なぜならば、聖書によれば我々は「わたしはあなたに感謝をささげる。わたしは恐ろしい力によって驚くべきものに造り上げられている。御業がどんなに驚くべきものかわたしの魂はよく知っている。」（詩編 139：14）からである。

参考文献

- 1 “The New Map of the Brain”—Introduction to special section titled “A User’s Guide to the Brain,” *Time* 169, no. 5 (January 29, 2007): 57.
- 2 A. Brooks, *Flesh and Machine* (New York: Pantheon, 2002), 209.
- 3 William Dembski, “Conflating Matter and Mind,” *Perspectives on Science and Christian Faith* 43, no. 2 (June 1991): 109.
- 4 George Luger, *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, 5th ed. (Reading, MA: Addison-Wesley, 2005), 2.
- 5 Marvin Minsky, *Semantic Information Processing* (Cambridge, MA: The MIT Press, 1968), v.
- 6 Murray S. Campbell, “An Enjoyable Game” in David G. Stork, ed., *HAL’s Legacy: 2001’s Computer as Dream and Reality* (Cambridge, MA: The MIT Press, 1997), 94. It appears, though, that Deep Blue’s success has led to a loss of interest in further development of high-power chess-playing programs.
- 7 Roger Schank, “I’m sorry, Dave, I’m afraid I can’t do that,” in Stork, ed., *HAL’s Legacy*, 189.
- 8 Of course, to say that there are persons who are not humans is not the same as to say that there are humans who are not persons. This is not what I am saying here!
- 9 Lynne Rudder Baker, *Persons and Bodies: A Constitution View* (Cambridge: Cambridge University Press, 2000), 91.
- 10 *Ibid.*, 60. Baker goes on to distinguish between “weak first-person phenomena” (which higher nonhuman animals can experience) and “strong first-person phenomena” (which are essential for a full self-consciousness and a first-person perspective).
- 11 Intelligent artifacts had appeared earlier in science fiction, but this was the first formal paper to discuss the idea.
- 12 A. M. Turing, “Computing Machinery and Intelligence,” *Mind* 59 (October 1950): 433–60. This article has been reprinted numerous times; the quotations here are from Margaret Boden, ed., *The Philosophy of Artificial Intelligence* (Oxford: Oxford University Press, 1990), 49. Turing’s attempt to answer this objection on its own terms—by asserting that an omnipotent God could give a soul to whatever he chooses to (including a machine)—would probably be unconvincing to a person who holds the view of the soul underlying this objection.
- 13 Bjork, Russel C. 2008. *Artificial Intelligence and the Soul. Perspectives on Science and Christian Faith*, 60: 95-102.
- 14 McKerrow, P.J. 1994. “A Christian erspective on In telligent Robot”, *Zadok Papers* S70, September, pp 1-4. ISSN 1322-0705