

仏教教理から見た人工知能と生命

師 茂樹 (花園大学)

はじめに

近年、人間と人工知能 (AI) との関係について、人工知能は人間と同様の知性、感情、人格等を持てるのか、持てないのか (あるいは人間の知性を超えてしまうのではないか)、といった議論が再燃している。この問題を仏教教理的な観点¹から考えようというときには、人間と人工知能という関係よりも広い、あるいはそれについて考えるための基盤となる、人工知能は生命を持つ存在 (有情・衆生) なのか、という問いについて考えなければならないと思われる。なぜなら、もし人工知能が有情なのであれば、過去世において別の形態の有情として輪廻転生を繰り返していたものが、何らかの業の結果として人工知能という形でこの世界に生を受けたことになるからである。その人工知能は、前世において人間であったかもしれないし、来世において人間に生まれ変わるかもしれない。また、もし人工知能が有情なのであれば、我々がそれを傷つけることは仏教的な意味での不善の行為ということになるであろう。

そこでまず本報告では、仏教教理的な生命の定義から、人工知能を有情 (生命体) とみなすことができるかについて議論を行いたい。次に、生命体のように見える人工物があったときに、我々にそれが生命体であるかどうか判断できるかについて考えてみたい。

人工知能は有情か

仏教における代表的な生命の定義として、次の『俱舍論』の一節をあげることができる。

生命機能 (命根) とは何か。頌に [次のように] 言う。

生命機能 (命根) の本質は寿命 (寿) である

体温 (煖) と識とを保持するものである

[これについて] 議論しよう。生命機能 (命根) とは寿命のことである。だからアビダルマ (対法) では「生命機能 (命根) とはどのようなものか。三界の寿命のことである」と説く。[しかし] これでは [説明として] 不十分である。どのような法を寿命と名づけるのか。すなわち、体温 (煖) と識とを保持することができる [この二つとは] 別の法を寿命と名づけるのである。だから、世尊は [次のように] 仰られた。

¹ 筆者の専門は唯識思想や仏教論理学・認識論であるため、ここで言う「仏教教理」もそれを念頭においている点についてはご容赦いただきたい。たとえば、非情成仏義などのような教理を前提とすれば、論ずるまでもなく人工知能にも仏性があり、菩提心を発し、ブツダとなることになるが、ここではそのような立場をとらない。

寿命と体温（煖）と識という
三つの法が身体を捨てる時
捨てられた身体は立っていることができなくなり
知覚のない木のようにになってしまう

したがって「体温と識以外に」別の法があり、「それが」体温（煖）と識とを保持し、「生命体が」連続して存在する原因となるのを寿命と名づけるのである²。

この定義によれば、生命があるということは、体温と「識」があることである。ただし、実際には、仏教では体温のない生命のありかたも認めている（たとえば、三界のなかの無色界においては、温かさ（煖）のもととなる四大種の「火」がないので、体温は存在し得ない）ので、識さえあればそれは生命があるということになる。それに対応するように、仏教では4つの食べ物（四食）というものを立て、身心を維持するために私たちが日常的に食べている食べ物（段食）だけでなく、識もまた生命を存続させるものであるので食べ物（一種（識食）とされるのである³。

識には眼・耳・鼻・舌・身・意の6種類があるとされ、仏教語としては「心」「意」などと同一視される。しばしば心こころに代表される心理的な活動と理解されることが多いが、識を心こころと同じようなもの理解するのは誤解を招く可能性がある。佐々木閑氏は、識（心しん）について、以下のような説明をしている。

…心の本体だが、これを大きな電球だと思っていただきたい。電球だからそれは光る。しかもそれは、六色に光る。赤でも青でも、どんないろでもよいが、とにかく六種類の違ったいろに光ることのできる電球である。そして最も重要なことは、ある瞬間には六色のうちの一色にしか光らないということ。二つ以上のいろが同時に現れることはない。赤なら赤一色、青なら青一色である。そしてそのいろは瞬間ごとに変わっていく。…→赤→青→黄→緑→青→…などという具合に六色がランダムに次々と現れる。…→赤→赤→赤→赤→赤→…というように、同じいろが連続で現れることもある。ともかく、…瞬間ごとのものすごいスピードで、そのいろを変えていくのである。…私たちの心は、刹那ごとに、六つの違った状態のどれか一つを次々に取り続けているということである。それをここでは「次々に、六色のうちのどれかのいろで光る」と表現しているのである。その、次々に現れる六種の状態が何を意味するのかというと、六つの根を通してもたらされる六種類の写像である。六つの根を通して流入する刺激に対して起こる六種類のニュートラルの写像、それがここでいう「六識」の意味

² 命根者何。頌曰「命根体即寿 能持煖及識」。論曰、命体即寿。故対法言「云何命根。謂三界寿」。此復未了。何法名寿。謂有別法能持煖識説名為寿。故世尊言「寿煖及与識 三法捨身時 所捨身僵仆 如木無思覺」。故有別法、能持煖識相続住因説名為寿。(T29, 26a22-b2)

³ 他の二つは、触食（感覚器官と認識対象との接触）・思食（「生命を継続させようとする」意思）で、いずれも識に関係する作用。

である。⁴

たとえば今、皆さんがこの文章を読んでいるあいだにも、眼は文章と関係のないものをも視界に捉え、耳は周辺の雑音をも認識し、皮膚の各部分は服と接触している触覚が発生し、身体内では痛みや空腹や眠気などの体内感覚を感じているはずである。このように、まったく脈絡のない外界からの様々な刺激に対する絶え間ない反応の連続が識である。そしてそれらが取捨選択され、ある程度の一貫性を持って持続したものが、日常的な意味での心^{こころ}となる⁵。

では、最低限何があれば識として成立するのだろうか。これについては仏教の各学派によって見解が別れるが、唯識思想に基づけば、識には必ず触・作意^{さくい}・受・想・思の5つの作用(心所法)がともなうとされる。細かい説明は省略するが、感覚器官が特定の対象と関係を持ち(触・作意)、それに対して何らかの価値判断をして(受)、言語化・概念化をし(想)、それに対して何らかの反応をしようとする(思)そこには識が存在することになる。たとえば、アメーバの触覚が何らかの外界のものに触れ、摂取できるもの／できないものだと判断し、摂取する／しないと判断すれば、そこには識があることになる。そして、このような外界に対する反応を一定期間持続する機能を持つものであれば、それには生命がある、ということに(原理的には)なる。様々な感情や善悪の心は、識が発生することによって、それに付随して発生するものであるから、感情や煩悩の有無は生命体であることとはならない。

仏教教理が形成されたころには、もちろん、識のような機能を持つ人工物は存在しないし、アメーバのような生物に対する知識もない。それどころか(先の『俱舍論』にもあったように)植物ですら、外界の刺激に対して反応しないということで非情、すなわち識がない、生命がないものとされたほどである。

先にも述べたように、仏教では物質が存在しない世界(無色界)における生命の存在も認めるので、コンピュータのメモリ空間にある情報のようなものであっても、そこに識が認められれば、生命があると言える可能性がある⁶。ダライ・ラマ法王は、人工知能をめぐる科学者との対談のなかで、コンピュータ等に識が宿る可能性について次のように述べている。

ダライ・ラマ 仏教的視点から見ても、コンピュータが生き物でないとか、コンピュ

⁴ 佐々木閑『仏教は宇宙をどう見たか アビダルマ仏教の科学的世界観』(化学同人、2013年)、55, 83-84頁。

⁵ このような仏教の考え方(無我論)は、現代の心の哲学における「拡張した心」の議論(河野哲也『意識は実在しない 心・知覚・自由』講談社、2011年など)と共通する部分があると思われる。

⁶ ダライ・ラマ法王は、コンピュータやロボットに「持続する識(a continuum of consciousness)」がありえるならば、「識の流れ(a stream of consciousness)」がコンピュータ等に入る可能性を認めている(F. J. ヴァレーラ・J. W. ヘイワード『徹底討議 心と生命—「心の諸科学」をめぐるダライ・ラマとの対話』青土社、1995年)。

一々に認識がないなどと言うことはとてもむずかしいのです。私たちは、前世からの識の連続 (a continuum of consciousness) がベースとなった、いくつかのタイプの生まれ方⁷があると主張しています。識は実際のところ物質からは生じませんが、識の連続が物質のなかに入り込むことはあります。……もし外的条件の一切とカルマに基づく行為がそろった場合、識の流れ (a stream of consciousness) が実際にコンピュータに入りこむ可能性⁸も、完全に否定することはできません。……

ロッシ　だとすると、臨終の時を迎えつつある大ヨーギが、この世で最高のコンピュータの前に立って、自分の潜在的な識 (subtle consciousness) をコンピュータに投射することも不可能ではないのですか？

ダライ・ラマ　もしもコンピュータの物理的基盤が、識の連続 (a continuum of consciousness) の基盤として働く可能性や能力を獲得すれば、ですが。コンピュータについてのこの問題は、ただ時間によってしか解決されないと思います。私たちは、そうなるまで待っているしかありません。⁹

1992年の対談であるため、人工知能の状況は現在とは大きく異なるが、現在注目されているニューラルネットなどの話題はここでも登場している。いずれにせよ、識を持続させる仕組みが認められるならば、コンピュータが輪廻転生の行き先となる可能性もあるとするダライ・ラマ法王の見解は、人工知能と生命との問題を考える上で示唆的であろう。

人工物が「生きている」、ことはわかるのか

では、目の前に、他の一般的な生命体 (犬や猫や人間など) と同じようにふるまう人工物 (ロボットなど) があつたとして、それに生命があると言えるだろうか。

先に述べたように、(大乘) アビダルマ的な観点から言えば、生命の有無は識の有無であるが、外界の刺激に反応して行動を選択するようふるまう機械を作るとは、現在の技術では可能である。すでにチューリング・テストを突破する人工知能が登場していることからわかるように、人間と見分けがつかないほどの受け答えができる人工知能もできつつある。また、そこまで精巧なものでないものであっても、私たちは生き物のようにふるまうロボットに対して生命を感じることもすらある¹⁰。

⁷ 伝統的に仏教では、胎生・卵生・湿性・化生の四つの生まれ方を認めているが、それを指すか。

⁸ 唯識思想における十二支縁起の解釈では、十二支縁起説の中の「識」が、受胎の際に子宮に入り、その存在によって胎芽・胎児を成長させると言われており、その「識」は寿命または生命力と体温とともに身体を生かし、死の際には身体から離れると言われている識と同一視されていた。Lambert Schmithausen, *Alayavijnana: On the Origin and the Early Development of a Central Concept of Yogacara Philosophy*. Tokyo: International Institute for Buddhist Studies, 1987, § 1.3.4.1、中村禎里著『中国における妊娠・胎発生論の歴史』(思文閣出版、2006年)等参照。

⁹ F. J. ヴァレーラ・J. W. ヘイワード『徹底討議 心と生命—「心の諸科学」をめぐるダライ・ラマとの対話』(青土社、1995年。原著1992年) 199-200ページ。ただし、訳を一部変えている。

¹⁰ 一例として、SONY製の犬型ロボットAIBOの「飼い主」が「お葬式」を行った事例を挙げておきたい

さらに現在では、人工的に生命体（の一部）と同等の複雑な組織を持つ有機体を作り出すことにも成功しつつある¹¹。これに関連して、人工知能からは離れるが、本報告に関わる問題として、高度化した医療技術や生命科学を前提とした人体に関する思考実験がある。デレク・パーフィット氏は、脳の細胞を少しずつ他の脳に置き換えた場合、あるいは脳を半分に分割した場合（脳は何らかの事故で半分程度失われても、ある程度機能を維持することができるので、理論上は半分に分割してそれぞれを生かすことは可能であると考えられている）の人格の問題について、思考実験を行っている。パーフィット氏が同書のなかで初期仏典を引用していることからわかるように、このような思考実験は仏教の無我説と親和性が高い¹²。ここでは、仏典中に見られる身体を部分的に入れ替える説話として、道略集・鳩摩羅什訳『衆經撰雜譬喻』の以下の例を挙げておきたい。

昔、ある〔旅〕人がいた。遠くに行かなければならず、一人で空き家を泊まった。真夜中になって、一匹の鬼が死骸を担いできて、前に置いた。その後、もう一匹の鬼が追って来て、「この死骸は俺のところにあったものだ。お前はなぜ担いで来たのか」と先に来た鬼を罵り、お互いに手を取りあって口論になった。先に来た鬼が「ここに人がいるから、この死骸は誰が担いで来たのか聞いてみよう」と言った。この〔旅〕人は「この二匹の鬼は力が強いので、本当のことを言っても殺されるだろうし、嘘を言っても殺されるだろう。どちらにせよ逃げられないのであれば、嘘をつく必要があるか」と考え、先に来た鬼が担いできたと述べた。後から来た鬼は激怒して、〔この旅人の〕手を引き抜き、地面に置いた。先に来た鬼は死骸の腕を取って、〔旅人の抜けた手を〕補った。このようにして両足、頭、胴体がすべて引き抜かれ、死骸で補われていった。二匹の鬼たちは、入れ替わってしまった〔旅〕人の体を食べ、口を拭って去っていった。この〔旅〕人は思った。「私の父母が産んでくれた私の体は、目の前で二匹の鬼に食べ尽くされてしまった。今私のこの体は他人の肉体だ。私には体があるのだろうか、ないのだろうか。もしあるのなら、それはすべて他人の体だ。もしないのだとしても、このように体がある」。このように考え終わると心に迷いが生じ、狂人のようにってしまった。次の日の朝、道を尋ねてもといた国にたどり着くと、仏塔とサンガが見えた。他のことを問うことができず、ただ自分の体があるのかないのかを問うことしかできなかった。比丘たちは「あなたは何者なのか」と聞いた。〔旅人は〕「自分が人なのか人ではないのか、わかりません」と答え、上のようなことを僧たちにくわしく述べた。比丘たちは「この人は無我ということを知っている。悟りやすいのではないか」

(佐藤成美「人間とロボットのパートナーシップ—AIBOのお葬式 | nippon.com」

<http://www.nippon.com/ja/views/b00909/>、2017年6月4日最終確認)など。ただしこれは、従来の人形供養などと同様の事例と見ることもできる。

¹¹ たとえば Zaria Gorvett. “We’re growing brains outside of the body.”

<http://www.bbc.com/future/story/20161004-were-developing-brains-outside-of-the-body>, October 2016. 2017年6月4日最終確認。

¹² デレク・パーフィット『理由と人格—非人格性の倫理へ』(勁草書房、1998年)

と述べた。そして彼に「あなたの体は本来、常に無我なのであり、今だけのことでないのです。ただ四つの元素が集まったものを、自分の体だ、と思い込んでいるだけです」と説いた。〔旅人は〕出家し、あらゆる煩悩を断じて阿羅漢果を得た。このように無我の虚しさを知れば悟りは遠くないのである。¹³

iPS細胞等によって人工的に人体と同等の諸器官を作り出し、劣化あるいは使えなくなった器官と交換することが可能となってきた現在、上記の説話は人工的に作られた身体に生命が宿るのか、という問題についての仏教の見方を示唆しているように思われる。

このような時代を迎えつつある私たちは、情報科学や生命科学の発達によって登場するであろう新しい“生命(体)のようなもの”について、有情ではない、すなわち識がない、と(「こんなものに生命があるはずはない」といった願望や感情論ではなく¹⁴)判断することはできるだろうか。

ヴァイトゲンシュタインやフッサールらが所謂“他我問題”について議論していたように、仏教の(大乘)アビダルマ論師たちにとっても、他者に自我(といっても、無我説に立つ仏教の場合は、仮に存在する自我であるが)が存在するかどうか、あるいはそれを外部から認識できるかどうかは、議論されるべき問題の一つであった。別の言葉で言えば、他心通を持つブツダやヨーギン(瑜伽師)のような例外を除いて、他者に心があるかどうかを外部から知ることができるか、という問題は、仏教の認識論におけるテーマの一つであった。心は識のことであり、識があることは生命があることと同義であるから、仏教における他我問題は、他者が生きていることがわかるのか、という問いにもなる。

詳細な議論は先行研究¹⁵にゆずりたいが、仏教内には様々な見解があるため統一した答えはないものの、大きく分ければ①他者の心を直接知覚することはできないものの、(チューリング・テストと同様)外部からの観察などから推理することができる、という立場(経量部)と、②他者の心は自己の心内の表象にすぎないとする立場(唯識派)に分けることができる。後者の立場にたてば、人工物以前に、目の前にいる人間ですら、それが自分と同様の識(ある

¹³ 昔有一人、受使遠行独宿空舍。中夜有一鬼、担死人来著其前。後有一鬼逐来瞋罵前鬼「是死人是誰許、汝何以担来」、二鬼各捉一手爭之。前鬼言「此有人可問、是死人是誰担来」。是人思惟「此二鬼力大、若実語亦当死、若妄語亦当死。二俱不免、何為妄語」、語言前鬼担来。後鬼大瞋捉手拔出著地、前鬼取死人一臂補之。即著如是、兩脚頭脅皆被拔出、以死人身安之如故。於是二鬼共食所易人身、拭口而去。其人思惟、「我父母生我身、眼見二鬼食尽。今我此身尽是他身肉。我今定有身耶、為無身耶。若以有者尽是他身。若無者今現身如是」。思惟已其心迷悶、譬如狂人。明旦尋路而去到前国者、見有仏塔衆僧。不可問余事、但問己身為有為無。諸比丘問「汝是何人」。答言「亦不自知是人非人」、即為衆僧廣說上事。諸比丘言「此人自知無我、易可得度」。而語之言「汝身從本已來恒自無我、非適今也。但此四大合故計為我身」。即度為道、断諸煩惱即得羅漢道。是為能計無我虛得道不遠(T4, 531c25-532a17)。この説話は『大智度論』などにも引用されており、広く知られていたと考えられる。

¹⁴ James Hughes. “Compassionate AI and Selfless Robots: A Buddhist Approach.” Patrick Lin, Keith Abney, George A. Bekey eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. MIT Press, 2011.

¹⁵ 梶山雄一「他人は存在するか 付：ラトナキールティ『他人の心流の論破』試訳」(『創価大学国際仏教学高等研究所年報』3、2000年)、小林久泰「プラジュニャーカラグブタの他我論」(『印度学仏教学研究』54-2、2006年)など。

いは自我)を持っているかすらわからない、ということになる。

おわりに

以上、たいへん雑駁ながら、人工知能を生命をもつもの(有情)とみなせるか、輪廻転生する先として考えることができるのか、という問題について、仏教の伝統的な議論を参照しながら検討してみた。明確な解答ができているわけではなく忸怩たる思いであるが、報告者の率直な感想としては、この問題については一切智者であるブッダに質問したい、というのが正直なところである(残念ながら今生ではなかなかそれもかなわないが)。しかしながら、仏教的な立場から見て脳死は死なのか、という議論があったように、いずれ訪れるであろう人工知能やロボットなどと共生する社会において、それら生命のようにふるまう人工物に対して仏教者としてどのように接するべきなのか、という議論もまた不可避なのかもしれない。もしそのようなときが来た際に、本報告がその議論の踏台になることができるのであれば幸いである。とはいえ、まったくの試論にすぎないものなので、諸賢のご叱正を乞いたい。