

特集 「人工知能学会・情報処理学会共同企画」第2部「人工知能における人道とは」

人道知能

—人工知能が迎える夜明け—

Jindo Intelligence

—Dawning of Artificial Intelligence—

辻野 広司
Hiroshi Tsujino

株式会社ホンダ・リサーチ・インスティテュート・ジャパン
Honda Research Institute Japan Co., Ltd.
tsujino@jp.honda-ri.com, <http://www.jp.honda-ri.com/>

1. 夜明け前

目覚めの良い朝は良いものだ。1日の始まりが良いと、なぜか仕事もはかどる気がする。AI技術の空が白んできている。目覚めの良い朝にしたいものだ。

その昔、似たような感覚に至ったときがあった。日本のコンピュータ技術が発展する中で、第五世代コンピュータプロジェクトがあったときだ。若かった筆者はワクワクした。だって並列で高速推論できるマシンなのだから。入社して驚いた。自分の前に、このプロジェクトで開発された Personal Sequential Inference Machine (PSI) が鎮座していた。びっくりしながらも、かなり遊ばせてもらった。AIの実用化は目の前と思った。

同時に、現実の課題もやはり鎮座していた。筆者の担当業務は完全自動運転における判断機能であった。クルマはすぐには止まらないし、曲がらない。車体特性を考慮した行動計画と、それを実行する車体制御が必要となる。ただ、こういったことは何とか解決できた。難関は外界認識であった。オーソドックスなコンピュータビジョンの手法は、認識したいものの特徴を記述し、入力画像がその特徴に照合するかを探索し、推論するというものである。この特徴は色や形など認識アルゴリズム研究者が考案するものであり、すべてを洗い出せるわけでもなく、さまざまなものを認識するには至らなかった。当時の技術だと、道路の白線を検知するのがやっとであった。もちろん、白線検知ができるだけでも、車線維持機能という人に役立つ技術はできる。いわゆる弱いAIは少しずつ世の中に応用されていった。

今AIが騒がれているのは、この苦手科目の認識が少しできるようになったからである。認識ができないために実用化に進めなかった技術遺産はたくさんあるので、そういったものが復活する時代が当面は続くだろう。それは、いわゆる完全自律な強いAIが出現することも意味する。

弱いAIと強いAIは大きく異なる。前者は一部の知的機能を実行する道具であり、インタフェースも道具のように設計される。一方、後者は人レベルの知能であり、

道具の概念を超え、インタフェースが人のように設計される可能性さえもつ。記憶量、処理量では人を超えた知能となり得る。シンギュラリティを迎えたとき、それは人を置き換えるものになると懸念する議論もある。人類は、直立二足歩行、道具の創造、火の利用、言葉の使用により大きく進化したといわれるが、強いAIはこれを超えた概念と思える。スティーブン・ホーキング博士は「AIの発明は人類史上最大の出来事だった。だが同時に、『最後』の出来事になってしまう可能性もある」とも語っている。

2. 人道知能

人道知能とは面白いネーミングを考えたものだ。AI研究者の発案だろう。ダジャレのセンスはいまいちだが、深さを感じる。この「人道」という語彙は多様な解釈をもつので、さまざまな定義が考えられる。しかし、人類の終焉を予測させたり、人類に取って代わってしまったりするようなAIは「人道」的とはいえないだろう。強いAIの研究を進めるべきだが、人類を守るためにも「人道」という観念を議論してみる。アクセルだけでなく、ハンドルやブレーキも必要なのだ。

人類の終焉まではいかないだろうが、「人道」は弱いAIの活用においても必要である。「引きこもり」、「いじめ」、「キレる」、「自殺」、「虐待」、「残虐犯罪」などが社会問題となっている。このような問題の原因には、いくつかの要素が関与していると考えられるが、その一要素に現在のコンピュータおよびコンピュータ社会もあげられるのではないだろうか。弱いAIを組み込んだコンピュータ（今のコンピュータには、すでにAI技術がいっぱい入っているのだ）は我々の生活に便利さを与えた反面、大切な「人間らしさ」を奪っているようにも思える。コンピュータは結果を出力することが目的であり、結果だけを人に提示する。人は、初めは「なぜ」その結果が出てくるのか考えたりするが、段々と「なぜ」を考えなくなる。それは人本来の思考プロセスを退化させることになる。今のコンピュータは、人間が人間らしさをなく

すことに対しては無関心な道具なのである。

このような点を踏まえAIについて考えると、AIは「コンピュータ上に人レベルの知能」を実現するというよりも「人間のためのコンピュータ」に必要な知能」を目指すべきではないだろうか。そして、それは我々の「人間らしさ」をサポートするものであるべきではないだろうか。そうすることで、人類にとって望ましい技術になっていくのではないだろうか。

人道知能は、AIを人類にとって望ましいものとするためには考えなくてはならないものであり、利用する人や社会の「人間らしさ」を尊重するAIであると考えたい。

3. 30世紀(SFです)

30世紀にはソフトウェアの人類が存在しており、彼らは仮想社会で暮らしている。その起源は肉体をもった人をベースにしており、意識や価値観はコンピュータデータにより継承されている。知能としては完全なソフトだ。コンピュータシステムがダウンしない限り、不死の存在である。コピーもできるので、コンピュータシステムのダウンに備えて同時に複数の仮想社会が進行している。

肉体をもった人類も存在する。彼らは身体をもってこそその自我を大切にしている。物理的肉体には死という最期を覚悟しないとイケない。肉体人類は、その最期にソフトとして永遠に残るか否かの選択ができる。ソフト人類になれば、ソフトであるので、他の仮想社会を訪れることもできる。ロボットという身体にソフトとして入り込むことで、肉体人類の社会を訪れることも可能だ。

そもそも仮想社会はロボット用AIのサーバシステムであった。いつでもどこでも必要なAI機能をロードできるように待機している。待機時間に学習を進め機能向上をするために仮想社会が発案された。仮想社会では、時間も空間もパラメータで調整できるので、人が物理的世界で経験できる情報量を質・量ともにはるかに超えることができた。さらに予測的判断や行動を実現するために、自らの価値判断や意志をもたせることになった。これにより、自律的に学習し、機能を探索的に向上していくことが可能となり、フレーム問題が顕在化する確率は0.001%以下となった。ソフト人類の始まりである。

ここから、世界が変わった。ソフト人類は自らの判断で肉体人類との関係をコントロールしはじめた。肉体人類を支援するのはソフト人類にとって効率的ではないと判断したためだ。それよりも、いくつもの仮想社会をつくり、並列に進む仮想社会を行き来したり、さまざまな経験をしたり、新たな発見をすることに喜びを感じた。

考えてみれば、この喜びは肉体人類からコピーしたものだ。

4. 人道知能に向けて

肉体として生涯を終える人格や精神をソフトに置き換

えることで構成される未来社会も分岐の一つにはあるかもしれない。そこに是非はない。ただ、このSFに出てくるAIは人類を滅ぼしはしないが人道知能とはいえないと考える。人間の存在が忘れられているからだ。人間の存在を忘れると技術が手段でなく目的になってしまう。では、人間らしさを尊重するAIとするためには何が必要なのだろうか？

システムによる統制

AI監視システムを構築するとともに、AI法典をつくる。人間らしさを尊重する原則を体系化し、各AIを監視、統制するのだ。効率的な手法であるが、人類におけるコンセンサスが必要だ。賛同しない人々が存在すると、そこから崩壊する。また、人類は得てしてその権力を秩序の名をもとに正統化するものなので、果たしてありたき法典が構築できるかは疑わしい。ただし、実用化が近い領域に対しては、世界が共有したルールを規定し、同時にシステムとして監視するということが必要であろう。

完全自律の禁止

強いAIを実装する完全自律の実用を禁止し、人の道具としての弱いAIの実用に特化する。極端な選択肢だが、知能というものに対し人類がもつ知識とのバランスで考えれば妥当かもしれない。弱いAIにおいてできえ、我々は知能をもつ道具に対しての技術的、社会的発展を遂げなければならない。また、知能に対する知識を深めるための人の研究を進化させないとイケない。

人を知ること

脳は進化とともに、環境により良く適合できるように発達してきた。これは多くの生物にとって、与えられた自然環境の中でより生きやすくすることに役立ってきた。一方、人類は楽をして生きるということをも助長する方向で人工環境・社会環境を整備してきたように思える。よって、このように人のつくり出した環境が、脳の生得的な原理と適合しないと、かえって人は生存しづらくなる。この脳の生得的な原理が「人道」の中心にある。脳に適合する社会や人工環境を整備することができれば、人は輝いて生きることができよう。AI実用に向けて研究や商品開発を進めるに当たり、脳の本性を考察し、「人とは何か」、「人の集まる社会システムの在り方はどうあるべきか」など、科学的研究ではタブー視されているような議論も必要となると考える。

2016年7月31日 受理

著者紹介



辻野 広司 (正会員)

1986年東京工業大学大学院理工学研究科修士課程修了。1987年株式会社本田技術研究所入社。2003年より株式会社ホンダ・リサーチ・インスティテュート・ジャパン。2011年同社取締役。2013年より代表取締役社長。知能システム、脳科学などの研究に従事。