

---

# 「人工知能」の研究とは？

植 村 渉

Wataru UEMURA

理工学部電子情報学科 助手

Research Associate, Department of Electronics and Informatics

---

## 「人工知能」とは

「人工知能」。この研究分野を聞いて、あなたは、どのような研究を想像しますか。ドラえもんや鉄腕アトムの実現、将棋やチェスといったボードゲームにおいて人間への勝利、行き先を伝えるだけで自動運転してくれる自動車、知能を持ったロボットの反乱。世間一般的には、このような「未来の夢」を「人工知能」として捉えていることが多いと思います。

ここでは「人工知能」の研究が、どのようなものであるのか、また、未来の夢の実現にむけてどれくらい進んでいるのかを、大雑把にですが紹介します。

私は小学生の時に、コンピュータに触る機会があり、そのときからプログラミングを行ってきました。そのため、情報工学は正規の教育課程として受講する以前から、独学で進めてきました。大学入学時に、残念ながら第一希望<sup>1)</sup>の情報工学科に落ち、応用物理学科に入学しました。その間も、独学で進めてきましたが、限界を感じ正規の教育を受けようと決意し、大学院で情報工学専攻として受験しなおしました。その際、どこの研究室に所属するか選ぶ必要がありました。つまり、どの分野を研究するか、決めなければなりませんでした。

せっかく情報工学を正規に受けて研究をするのな

ら、独学でできない分野をしたい。こう考え、それぞれの分野の研究がどのようなものか調べてみました。ここで、冒頭の質問に戻ります。当時の私は、情報工学の世界を、外から眺めていました。私の考えていた「人工知能」の世界は、冒頭のような「自律したロボット」を目指す研究でした。そういう知能が実現するかというと、「現代の錬金術」と呼ばれるくらい難しいです。でも、難しいからこそ、研究を進める価値があるはずですし、進めているだろうと考えていました。しかし調べてみると、人工知能関係の学会の論文誌に投稿されているのは、(私にとっての)「自律した知能」ではありませんでした。それらの論文に書かれている技術を否定するつもりはありません。技術として、適用分野があり、ちゃんと世の中の役に立っています。でも、私のイメージしていた「人工知能」とは異なったのです。

## 「人工知能」がとった行動の責任

「人工知能」の研究を続けてきた今、振り返ってみますと、やはり「自律した知能」が数年で実現するのは無理だと思います。無理であるからこそ「今できる知能」を研究してきた結果が、論文誌に書かれている技術であったのです。例えば、人工知能を題材とした映画などでは、知能が自我を持ち、反乱

を起こすことがあります。そのときに、誰が責任を持てるのでしょうか。そのロボットを作った人に、その後の成長を含めて責任を負う必要があるのでしょうか。そのロボットを使っている人に、そういう成長をさせた責任があるのでしょうか。この議論は、もはや工学の域を出ます。

責任が生じないためには、どういう研究をすれば良いのでしょうか。そのひとつの方法として、「提案」があります。計算機が人間に、「こういうのは、どうですか?」と提案し、最後の決定権は人間に渡します。このことにより、最終的な責任は、人間に移ります。

最近の車に搭載されている追突軽減ブレーキ<sup>2)</sup>などは、いい例です。前走車に近づくと、シートベルトを引きブザーを鳴らして運転手に知らせ、ブレーキをかけます。しかし、そのままでは、ぶつかるようになっています。ぶつからないブレーキをかけることも可能でしょうが、前述のとおり、最後の決定権は人間が持つようになっています。自動的にハンドルを切って車庫入れをする車<sup>3)</sup>もあります。これも、ハンドルは車が操作しますが、アクセルは人間です。万が一、進路上に障害物がある場合、車を動かしたのは「人間」になります。いずれの車も、今までの車と同様に、人間が適切に操作していれば、事故が防げるわけです。こういう「提案する知能」<sup>4)</sup>は、実用化の問題を解決する意味で、重要です。しかし、能動的な「自律した知能」と比べると、受動的な知能です。私は、能動的な知能を目指しました。

## 閑話休題

本当に「自律した知能」が実現したらどうなるのでしょうか。権利のない労働媒体である奴隷ができるのか、それとも、人間と対等の立場となり、労働力としてはあまり意味を持たない存在となるのでしょうか。

四次元ポケットを持たないドラえもんを想像してみてください。単なるおせっかいなロボットです。

しかも、よく、失敗します。本当に、そんなロボットが欲しいのかと言われると、ちょっと躊躇します。自律したロボットの正体は、案外そんな程度のものなのかもしれません。

## 生体の知能

人工的な「知能」の実現を考える際、対として生体の「知能」がどうなっているかという解析も必要です。よく用いられる例ですが、「空を飛びたい」という夢に対して、鳥の真似をしても飛ぶことはできませんでした。理屈を考え、計算の結果できあがったものは、羽ばたくことのない、重たい鉄の塊である「飛行機」でした。つまり、生体の構造を追及するよりも、発想の転換で、新たな構造を考えるほうが、実現する場合もあるということです。ですので、生体の「知能」の真似をしても「人工知能」の実現は保証されません。しかし、知能に近づくために生体の「知能」を解明することは、決して無駄にはならないでしょう。

## 強化学習

私が着目した研究は、「強化学習」と呼ばれる分野です。強化学習とは、学習者であるエージェントが、試行錯誤を繰り返し、目標状態(=ゴール)に到達した経験を基に、今までの行動選択が正しかったかどうかを判断し、強化する方法です。例えば、ロボットの入力センサーから得られる離散情報を状態とし、その状態において選択できる行動群の中から一つを選択し(=状態における行動をルールと言います)、実行することでロボットが移動したりして状態が変化するという枠組みです。このときの、各状態において、どのルールを選択するかという選択方法(=政策もしくは方策と言います)が、賢さを決めます。

現実世界での動作を考えると、行動を実行した際に、必ずしも同一の状態変化を期待できない場合も

あります。その場合は、確率的な状態遷移となります。簡単なモデルとしては、状態遷移が決定的な場合となります。将棋やオセロゲームといったように、盤の駒の配置を状態とし、次にどの手を指すかが行動となる場合です。ある盤の状態で、決まった手を指せば、次の盤の状態は必ず同一です。このような簡単な場合でも、問題空間は大きくなり、最善手を知ることができない問題となります。

## Profit Sharing法

強化学習の中でも、Profit Sharing法と呼ばれる学習方法に魅力を感じました。硬い説明としては、目標状態に至るまでの行動選択をエピソードと呼び、目標状態に到達したときに得られる報酬 $r$ を、それまでのルールに割り振ることで選択確率を高め、ゴールに到達できるルールを強く選択する(=学習する)方法です。

実は、この一連の流れは「餌付け」の作業に類似しています。ねずみが迷路を走行し、ゴールに着いた時に餌を与えます。すると、ねずみはその経路を覚え、何度か試し学習することで、素早くゴールに到達できるようになります。そう、Profit Sharing法は、生体の知的作業にかなり類似しています。

幸いなことに、私が最初に興味を持った強化学習がProfit Sharing法でした。しかし、強化学習の代表的な手法は、Q-Learningと呼ばれる手法であり、Profit Sharing法とはかなり本質の異なる方法でした。Q-Learningでは、将来的に獲得できる期待報酬値を各ルールの価値として定義し、何度も思考することで、各ルールの価値が期待値に近づくように更新をします。簡単に分類すると、強化学習の王道のQ-Learningなどの手法では、「将来、どれだけの報酬が得られるか」を求めようとし、Profit Sharing法では、「今まで、どれだけの報酬が得られたか」を基に行動を選択する方法となります。Q-Learningでは、将来の獲得報酬値を得ようとするので、条件がそろえば、最適解を学習することができます。そ

れに対して、Profit Sharing法では、過去の経験にのみ目を向けているため、獲得した解の最適性は保証されません。ただし、今までの経験から、よい行動選択を考えますので、確実性が高く、学習の立ち上がりが早いという特徴があります。

## 人間らしい特徴

さて、ここで、Profit Sharing法の持つ学習の性質を「人間らしさ」という観点で見えます。今現在、私が気づいている「人間らしさ」は、三点あります。

まず、一点目は最適解の学習が保証されない点です。最適性がある学習方法が存在する中で、最適性がないという特徴は、工学的にはデメリットです。しかし、「自律した人工知能」を考えると、この特徴は、かなり興味深いものとなります。計算機は、指示されたとおりにしか実行しません。融通が利かないかわり、正確な解をだす能力を持っています。このような正確無比な計算機において、解を導出する際に、最適解よりも、そこそこの解を導出するというのは、聞いたことがありません<sup>5)</sup>。計算機が、「曖昧さ」を持つことができるのです。どことなく、「人工知能」の実現の糸口が見えてきませんか？

二点目の「人間らしさ」は、「素早い学習」です。先ほどの最適解の学習が保証されない点と関連しますが、最適性を保証するためには、全パターンを知る必要があります。そのため、組み合わせが大きくなり状態空間が爆発すると、実時間では最適な解を計算できなくなります。しかし、Profit Sharing法では、過去の経験に基づいて行動を選択します。今までに報酬を獲得できた行動を優先して選択するため、安定して報酬を獲得し続けることができます。結果として、学習の立ち上がりが早くなります。

そして、最近着目している三つ目の特徴は「報酬の期待値に合わない行動選択をする」ことです。この特徴も、工学的にはデメリットです。期待値に合わない行動選択をするため、性能評価的には悪い結

果となります。原因は、Profit Sharing法では、報酬獲得時にのみ学習作業をするため、小さい確率で大きな報酬を得られる行動を学習初期に経験すると、その後の行動選択に大きな影響を与えます。例えば、最初の学習時に、 $p$ の確率で起こる状態遷移を経験したとします。Profit Sharing法では、過去の経験にのみ目を向けるため、自分の中では、 $1/1$ の確率でその事象が起こっています。 $p$ が $1/10$ であろうと、 $1/100$ であろうと、起こった出来事であるため、違いがわからないのです。実は人間も、同様の計算ミスをすることがあります。競馬やパチンコといったギャンブルは、胴元が儲かっていることより、明らかに参加者の期待値はマイナスです。特に宝くじに関しては、参加者の実力の関係する要因がなく、全員平等に期待値がマイナスのゲームです。期待値が明らかにマイナスとわかっていながら、購入します。数学的には、説明がつかない行動選択です。「曖昧さ」を持つ「人工知能」の実現を考えると、この特徴は面白い特徴だと思います。しかし、実用化においては、活かすのが難しく、不要な特徴です。

以上のように、Profit Sharing法には、現段階で私が気づいているだけでも、三点の人間らしい曖昧さを持っています。そして残念ながら、工学的にはデメリットです。おそらくこれ以外にも、人間らしさを持っているとらんでいます。今後は、それらを見つけ出すとともに、「人工知能」の実現の糸口として、提案していくことを考えています。

## まとめ

今回は、「人工知能」という研究が、どのようなものであるのか、そして、人工知能の中でもProfit Sharing法が人工知能の実現に対してどのような特

徴を持っているのかを、大雑把にですが、紹介しました。文章中心で、数式やアルゴリズムは一切使いませんでした。実際の研究では、具体的にどのようなアルゴリズムで学習が進むのかを議論したり、行動選択に用いる評価値がどのように更新され、どのような値になるかを数学的に議論したりしています。また、理論を進める際には、理想的な問題環境の条件から、より現実的な厳しい条件へと拡張していきます。最終的には「学習に必要な情報が足りない中での学習」という条件になります。普通に考えれば、そのような条件下では、正しい判断ができなくて当然です。しかし、本当にそうでしょうか。「ぼやけた人物画像から、元の画像を復元する」という研究<sup>6)</sup>があります。この研究では、画像復元のために必要な情報が足りない条件で画像復元を試みます。多数のサンプル画像を用意し、統計的に扱おうと、十分な画像が復元できたそうです。「条件が整っていないからできない」とあきらめるのではなく、その条件下でできることを考え研究すれば、できる(かもしれない)のです。

人工知能の研究では、発想の転換が必要な時期にさしかかっています。みなさんも「知能」について考えてみませんか。

## 注

- 1) 著者出身の阪市大の工学部の入試では、複数の学科を希望することができました。
- 2) ホンダのインスパイアに搭載 (CMS: Collision Mitigation brake System)。
- 3) トヨタのプリウスに搭載。
- 4) なお、これら自動車の例を「提案する知能」と解釈するのは、公式ではなく、私の独断です。
- 5) 山登り法なども、最適性を持たず、局所解へ収束する解法です。しかし、知的な作業の結果ではないので、私の中では単純に局所解にたどり着く手法は、「曖昧さ」を持つ手法に属していません。
- 6) 第三回情報科学技術フォーラムFIT2004, 特別講演:「計算機視覚の研究-その仕組みと今後-」, 金出武雄