

近未来金融システム創造プログラム第9回講義レポート

第9回目となる本日は、理化学研究所脳科学総合研究センター特別顧問の甘利俊一氏から「機械学習:人工知能の歴史とこれから」というお題で講義が行われた。常に最先端のテクノロジーを取り込み、拡大と多様化を果たしてきた金融はこれからどこへ向かって行くのか?「金融と技術」を語る上では切り離せない、人工知能の歴史と未来について数理脳科学の観点から語っていただいた。

生命の仕組み

「人口知能は人間の脳を超えるのか?」という問いを議論するためには、人工知能の仕組みや歴史、現在の社会との関係性を理解する必要がある。その初期段階として、人間の脳と人工知能の類似点・相違点について考えることが重要だ。

人間の脳を理解するためには138億年前のビッグバンまで遡る必要がある。ビッグバンは物質やエネルギー、時間、空間などをもたらした。生命も物質の一種だが、情報に基づいて自分と似た物質を複製し次世代につなぐ。このような自己増殖と進化が他の物質(単なる結晶など)との大きな違いだ。しかし、生命はその情報を物質の上にエンコードしているため情報に“揺らぎ”が生じる。この“揺らぎ”が外部環境に対してより生存性の強い生命を誕生させ、古典的な物理法則では説明不可能な進化の法則をもたらした。すなわち、生命は情報と物質の二つの基盤からなる神経細胞によって意識や思考を生み出したのだ。その結果、人類はコミュニケーションや言語を用いて文明や社会を築いた。

人工知能の歴史

コンピュータの起源は、1930年にイギリスの数学者が数学モデルから考案した万能性を有するチューリングマシンである。コンピュータが製品として普及し始めた50年代には、単なる計算だけではなく、チェスで人間を負かすようなより知的な推論を求めるようになった。このように、記号と論理のみで人工知能の開発を試みたのが第一次人工知能ブームである。一方、同時期には、経験と学習によって知的機能を獲得する人間を参考とし、機械を学習する普遍計算機構として捉える考え方も広まっていた。しかし、どちらも当時の技術力では実現することはできず、60年代後半から70年代にかけてのAI暗黒期を迎えることになった。

1970年代からの第二次人工知能ブームは、専門家のように大量の知識を蓄えて推論を行う、エキスパートシステムを作ろうという流れから始まった。ソフトウェア開発者が設定した手続きに直接従うのではなく、大量に蓄えた知識とif-thenルール(「もしこうなったら、こうする」という実行計画)によって複雑な問題の解決を試みた。実際に、医者が行う診断を機械で再現したが、その診断の質はそれほど悪くなかったものの費用対効果の点から実用化には至らなかった。

2010年からの第三次人工知能ブームの火付け役になったのが深層学習・ディープラーニングだ。この背景には、データの処理速度の向上をはじめとするコンピューターの技術的進歩が存在した。2010年代にはコンピューターの画像認識精度が人間の精度を上回り、2016年にはグーグルの開発したAlphaGoが囲碁で韓国のイ・セドル九段を4勝1敗で破ったことで機械学習が再び注目を浴びるようになった。

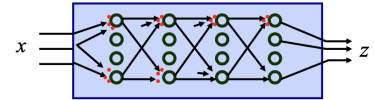
深層学習とは

深層学習は、神経細胞(ニューロン)のつながりを数理モデル化した「学習する層状の神経回路網(ニューラルネットワーク)」に情報を入力し、それぞれの神経細胞で計算を繰り返して答えを出す仕組みとなっている。一般的な深層学習には、学習データをシャッフルした上でそれらの中からランダムに1つを取り出し、誤差を計測してパラメーターの更新を行う、確率勾配降下学習法が採用されている。

ここでいう「学習」とは、正解の分かっている学習データを層状学習回路網(ニューラルネットワーク)に入れ、正解になるようにニューロンのパラメーター(重み)を調節していくことを指す。既に分かっている正解を目標値とした場合、ある学習データを情報として入力した時に生じる出力値と目標値との差を誤差とする。この誤差パラメーターを最小にする工程が「学習」である。

実際には、誤差としてパラメーターを含む「コスト関数」と呼ばれるものが定義される。構成されたニューラルネットワークの出力側から入力層までのニューロンの接続を見ていくと、コスト関数は最終的に多数の複合方程式として表現することができる。つまり、「学習」とはコスト関数が最小になるようにパラメータを修正していくことを意味する。

層状学習回路網
multilayer perceptron



パーセプトロン Perceptron
バックプロパゲーション Backpropagation

$$L(x, W) = |y - g(x, W)|^2$$
$$w \rightarrow w + \Delta w, \quad \Delta w = -c \frac{\delta L(x, W)}{\delta W}$$

人工知能と人間

人間が意識を獲得した理由は、共同生活・社会生活を送る中で、自分の意思を相手に伝える必要があったからだ。脳が一度決めたことをさらに吟味して自分の決定の合理化を行う一連の行動を意識と見ることができる。しかし、現段階の人工知能の行う機械学習は、入力されたデータを処理する中でパラメーターの調節によって、合理的且つ最適な決定を行うだけであり、「心や意識」さえも神経活動の調整によって合理的に導かれた一連のプロセスに過ぎない。

局所的に見れば人工知能の処理能力は人間の知能を既に超えており、シンギュラリティーが間もないという意見も存在する。このような議論の中で注目されるのが、ロボットに「心や意識」を与えるべきか？また与えることは可能か？という問いだ。心は合理的な判断を常に促すわけではないため、一見不条理なものだとも考えられる。だが、現在の技術力を応用すれば、人間が感情移入できるほどのロボットを制作することはできるだろう。しかし、ロボット自体が真に喜怒哀楽を持っているわけではなく、あくまで合理的な存在であることを忘れてはならない。

人工知能の発展は止まることはない。人間が今まで進化の過程で成長してきたように、人工知能も発展していく。人間は、堅実な制御システムの範囲内で無駄な仕事を人工知能に代替し、より豊かな社会や生きがいのある仕事を追求すべきだ。このように、一人一人の人間が生を全うすることで望ましい社会が実現されるだろう。

Q&A

Q, プレディクションとポストディクションの違いを具体的に教えていただきたいです。

A, プレディクションは、様々な可能性の中から計算によって一つを選択する操作。ポストディクションは、その時に考慮しなかった可能性の部分にバイアスをかけて考え直すことである。人間はある意思決定をしたときに後からその選択を正当化する。多くの場合は自分の決定を他の理論によって補強するのだ。

Q, 「長い文明の中で築かれてきた人間の感情移入」をデータでロボットに学習させた場合、ロボットは人間と同じように感情移入するのか？

A, 人間の感情は、個人の問題ではなく、社会の一部としての人間、個人と個人の関係性によって生じたと推測される。人間が生きて来た環境では、必ずしも合理的な判断だけが正解ではなかった。そのような複雑な環境の下では、感情を媒介としてコミュニケーションを行うことが最適だったのかもしれない。

戦略を合理的に考えるロボットにサッカーをさせた場合、チームメイト間でそのような感情のやり取りが行われるだろうか？もしかしたら、ロボットにも同じことが起こり得るかもしれない。