

書評

牧野貴樹, 澁谷長史, 白川真一 編, 浅田稔, 麻生英樹, 荒井幸代, 飯間 等, 伊藤 真, 大倉和博, 黒江康明, 杉本徳和, 坪井祐太, 銅谷賢治, 前田新一, 松井藤五郎, 南 泰浩, 宮崎和光, 目黒豊美, 森村哲郎, 森本 淳, 保田俊行, 吉本潤一郎: これからの強化学習, pp. 313, 森北出版 (2016)

最近のAIブームのけん引役の一つとなっているのは、AI囲碁の話題であることは疑いもないであろう。GoogleのAlphaGoが世界トップクラスの人間のプロ棋士に勝利してから、一気に深層学習 (Deep Learning) という言葉が広く知られるようになった。筆者のところに来る企業からの相談内容においても、学生の興味の対象としても、深層学習がこの1~2年でその多くを占めるようになった。そのように深層学習に焦点が当たることが多いが、AlphaGoの一つの特徴は自己強化学習にある。通常、深層学習を含むニューラルネットワーク系の学習の多くは教師あり学習に位置付けられ、与えられた教師情報に基づいてネットワーク内の重みを学習 (最適化) する。囲碁の問題では、過去の対戦を表す棋譜が教師情報となる。しかし、囲碁の盤面から得られる膨大な状態の可能な遷移すべての組合せに比べて、既知の棋譜はそのごく一部にしか相当せず、しかも人間同士の対戦結果から学んでも人間を超えることは難しい。そのため、AlphaGoでは、これまでに学習した行動規則をもつ仮想的な対戦相手を用意し、実際に対戦シミュレーションを重ねることで、さらに学習を進めるのである。

そのような強化学習とは、環境の中を自律的に行動するエージェントがその行動の結果もたらされる報酬を手掛かりに、最適な行動規則を学習する機械学習手法である。これは、何も特別なものではなく、我々人間も日々同じような学習をしている。例えば、子供は何かをして褒められるとそれを一生懸命するようになり、逆に叱られるとその行為をあまりしなくなることがある。子供に限ったことではないかもしれないが、この「褒められる」、「叱られる」というのが環境から与えられる報酬であり、最終的により多くの報酬を得るために、さまざまな行動に対する結果から、試行錯誤的に自らの行動規則を学んでいくことは、まさに強化学習である。

本書の第1章では、強化学習の基礎概念とその基本的な数理が丁寧に説明されている。具体的な例を用いた「エージェント」、「環境」、「状態」、「行動」、「報酬」、「収益」、「遅延報酬」などの基本概念とその関係性の説明は分かりやすい。そして、環境から受け取る報酬を最大化するためにエージェントが取るべき行動を学習する基本的なアルゴリズムが、数式の導出も含めて丁寧に述べられている。基礎的な参考文献も網羅されており、強化学習をこれから学ぶ者にとって良い教科書となるだろう。

一方、第2章は第1章で述べられた基本概念をベースとして最新の理論が紹介されている。この分野の第一線で

活躍されている研究者による解説は、読み応えのある内容である。だからといって必要以上に難しく論じられているわけではなく、こちらでも事例ベースでわかりやすく説明されている。統計的学習の観点から強化学習はどう論じることができるのか、強化学習の枠組みを設計するうえで報酬をどう設定すればよいのか、多数のエージェントが協調して一つのタスクを達成する場合にはどうすればよいのか、行動を学習するための試行錯誤の反復を軽減するにはどうすればよいのか、リスクを抑えつつ報酬を最適化するにはどうすればよいのか、など、これから強化学習の応用を考えるうえで検討すべき多くの事柄に気付かされる内容であり、それらに対する最新の情報を得ることができる。強化学習に関するより発展的な話題について知りたい場合、最初の調査対象としては丁度良いのではないかと思う。

第3章は実際の応用事例の紹介となる。第2章とは独立に読めるため、第1章で基本概念について学んだうえでこの章を読めば、実際の応用事例に強化学習の枠組みをどのように適用すればよいかを知ることができる。例えば、最も想像しやすい人工的なエージェントはロボットかもしれない。単体、もしくは複数のロボットがあるタスクを遂行するために試行錯誤的に最適な行動規則を学ぶことも強化学習で実現できる。また、そのような物理的な身体性を伴わないソフトウェアエージェントが人と自然に会話するための自然言語理解や対話処理にも強化学習は利用できる。それらへの強化学習の適用事例に加え、冒頭でも述べた囲碁への応用も解説されている。AlphaGoや、最近、インターネット上で覆面ソフトとして騒がれたその後継システムの強さの仕組みを知りたい方は、その解説を読むだけでも価値はあるだろう。

さて、子供が「褒められる」、「叱られる」という報酬のもとで自らの行動を最適化するという例を先に述べた。そのようなヒトの認知行動を脳科学的な視点でモデル化する際の道具立てとして強化学習を利用する取組みが、最後の第4章で述べられている。強化学習に対する基礎的な理解があれば十分に読める内容であり、認知科学に興味のある方には興味深いものと思われる。

この分野で活躍されている多数の研究者の知の結晶ともいえる本書は、これから強化学習の基礎を学びたい人の入門書としても、実務に機械学習を用いたアプリケーションの導入を考えている人の参考書としても、必要かつ十分な話題が網羅されており、そういう意味で本書はお薦めの良書である。ぜひ一読いただきたい。

[大原 剛三 (青山学院大学)]