

# ディープラーニングを用いた コンピュータ囲碁

～ Alpha Go の技術と展望～

応  
専



伊藤 毅志 村松 正和（電気通信大学）

## Alpha Go の衝撃

2016年1月28日、日本時間AM 3:00に、そのニュースは発表された。Googleの研究グループがNatureに“Mastering the Game of Go with Deep Neural Networks and Tree Search”という論文を発表した<sup>1)</sup>。それによると、ディープラーニングと強化学習を用いた手法で囲碁の局面を評価する新しい手法を確立し、その手法と従来のモンテカルロ木探索（MCTS: Monte-Carlo Tree Search）の手法を組み合わせることで、既存のプログラムに99.8%勝つばかりか、囲碁のヨーロッパチャンピオンのプロ棋士に5連勝したというものであった。

コンピュータ囲碁は、2006年頃に発表されたMCTSの手法で、近年めざましく進歩した。2015年3月に開催されたプロ棋士とコンピュータ囲碁の置碁公式戦である電聖戦では、第25世本因坊治勲を相手に、4子では勝利したものの3子では力の差を見せつけられており、プロ棋士とは3子以上の差はあると思われてきた。互先でプロ棋士に勝利するには、まだ10年程度は要するのではないかと考えられてきた。

チェスや将棋と比べ、コンピュータにとって囲碁は、盤面の広さによる探索の困難さもさることながら、局面の評価関数（盤面の優劣を数値化する関数）を作ることが絶望的に難しいとされてきた。囲碁では石1つ1つには意味がなく、それが連なって意味を構成していく。石の繋がり方は千差万別で、石1個の配置が違って意味はガラリと異なることがある。人間はこのような石の繋がりを、石の「強さ」

や「厚み」といった言葉で表現するが、このような感覚的な局面の理解はコンピュータには困難であると考えられてきた。

しかし、今回のGoogleの論文では、ディープラーニングを用いることで、局面を評価できるネットワークを構築したというのだ。これが、コンピュータ囲碁研究者・開発者を驚かせた。

## 予兆と Alpha Go が実現したもの

実は2014年から2015年にかけて、ディープラーニングをコンピュータ囲碁に応用した研究論文が立て続けに発表されていて、その予兆はあった<sup>2)～4)</sup>。これらの研究は、プロ棋士の棋譜データを教師データとするディープラーニングを用いることで、プロ棋士の手を予測するシステムを作るというものであった。それによると、これらのシステムが選んだ手は、プロ棋士の選んだ手と非常によく一致して、一致率が57%にも上るものもあった。これらの研究はそれまでの予測器の精度を大きく凌ぐものであり、ディープラーニングのコンピュータ囲碁への応用を期待させるものであった。

しかし、Alpha Goはコンピュータ囲碁研究者の予想を上回る成果を含んでいた。Alpha Goは、以下の3つの段階を経て構築されている。

### <第1段階>

13層のDeep Convolutional Neural Network (DCNN) を使い、インターネット囲碁道場KGSの6段～9段の棋譜16万局、約3,000万局面を用いて

教師あり学習 (Supervised Learning) を行い、次の一手の予測器を作成するフェーズ。50 個の GPU で 3 週間かけて学習を行い、これによって予測率 57% の高確率の予測器を実現した。この DCNN を SL Policy Network と呼ぶ。

### < 第 2 段階 >

上述の SL Policy Network を初期値とし、「ゲームの勝利」を報酬として自己対戦による強化学習 (Reinforcement Learning) を行うフェーズ。第 1 段階のものと対戦させた場合に 80% 以上の勝率を叩き出す DCNN を実現した。この DCNN を RL Policy Network と呼ぶ。この段階は 1 日で終了する。

### < 第 3 段階 >

自動生成したデータから、局面の勝率を計算する DCNN, Value Network を作成するフェーズ。まず、「途中まで SL Policy で自己対局させ、一手だけランダムに打ち、その後は RL Policy に打たせる」ことを繰り返し、3,000 万局面のデータを自動生成する。ランダムに打った直後の局面と勝敗がデータとなる。このデータから教師あり学習により、RL Policy Network によって局面の勝率を計算する Value Network を実現した。この学習には 50 個の GPU を用いて 1 週間で費やしている。

このようにして作成された SL Policy Network と Value Network, それに従来からの技術であるシミュレーションベースの MCTS を組み合わせて Alpha Go は構成されている。第 2 段階で生成された RL Policy は、第 3 段階のデータを生成するのに使用されるだけで、最終的には使われていないところが興味深い。

第 1 段階までは、すでに発表された技術であり、いくつかのコンピュータプログラムでも導入は進められていた。

第 2 段階以降が Alpha Go の新しい技術である。これによって、プロ棋士に匹敵する打ち手を実現していると思われる。特に Value Network は、不可能と言われていた囲碁の盤面評価関数を DCNN により実現したものであり、大きな衝撃である。

システム全体を構築するのに 50 個の GPU と 1 カ

月という期間を要すること、論文<sup>1)</sup>の筆者が 20 名にのぼること、また、プロと対戦した Alpha Go が 1,202 個の CPU と 176 個の GPU を用いていることは、いずれもこの研究が Google という大企業の豊富なリソースに裏打ちされたものであることを示している。

## Alpha Go の強さと今後の展望

Alpha Go の実力について、論文の中では、既存の有力プログラム (Crazy Stone, Zen, Pachi, Fuego, GNU Go) を相手に 495 局中 494 局 (勝率 99.8%) で勝利したとある。これをもって、Zen や Crazy Stone に対して、1,000 程度のレーティングの差があると論じているが、対戦したのは、市販版であり、クラスタ版と純粹に比較するのは、難しい。

また、ヨーロッパチャンピオンの Fan Hui (樊麾) 二段と 5 戦 5 勝し、その棋譜も掲載されている。実際には 10 局打って 8 勝 2 敗だったが、あらかじめ決めておいた「公式対局」5 局においては全部 Alpha Go が勝ったということである。「非公式対局」である 5 局分の棋譜は公開されていない。対局数が少ないこと、棋譜もやや淡白な内容であることなどから、プロ棋士の Alpha Go に対する評価も割れている。

3 月 8 日から 15 日にかけて、トッププロの 1 人である李世乭九段との対戦が予定されている。5 番勝負で賞金は 100 万米ドルのビッグイベントである。囲碁ファンには常識であるが、李世乭九段はまさしく世界のトッププロである。もし、ここで李世乭九段が負けるようだと、たしかにコンピュータが人間を超えたと認めざるを得ない。

しかし少なくとも、論文に書かれている技術を読む限り、これまで MCTS の問題点であった複雑な攻め合いの理解や、両コウを含むような難解な局面における問題が解決されているとは思われない。もちろん Fan Hui 二段との対局のように弱点を見せる暇なく序中盤で圧倒してしまう可能性もあるが、競った相手に難解な攻め合いに持ち込まれると、意外な

弱点が露呈する可能性がある<sup>5)</sup>。

人間対コンピュータの対局では、対局の公平性が議論されることが多い。公開されていない棋譜の中には、Alpha Go が負けた対局もある訳で、こういった対局も含め、李世乭九段に事前に十分な情報提供がなされ、可能な限り公平な対局が実現されることを願いたい。いずれにしても3月の対戦は大きな注目を集めることだろう。

電気通信大学では、この対戦の直後、UEC 杯コンピュータ囲碁大会が3月19、20日に、その優勝、準優勝プログラムが、小林光一名誉棋聖と対戦する電聖戦が3月23日に開催される。UEC 杯には、Facebook を含め、Crazy Stone, Zen, Dolbaram などの有力プログラムも参戦する。これらのプログラムも一部ディープラーニングを導入して、成果を見せつつあり、にわかにコンピュータ囲碁界は、X デーに向け加速しているように見える。

コンピュータ囲碁の根本的な問題点がこれから露呈してくるのか、それともその弱点を見せずに人間トップを押し切ってしまうのか、目が離せない。

参考文献

- 1) Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D. : Mastering the Game of Go with Deep Neural Networks and Tree Search, Nature 529, pp.445-446 (2016).
- 2) Clark, C. and Storkey, A. : Teaching Deep Convolutional Neural Networks to Play Go, arXiv:1412.3409 (2014).
- 3) Maddison, C. J., Huang, A., Sutskever, I. and Silver, D. : Move Evaluation in Go Using Deep Convolutional Neural Networks, arXiv:1412.6564 (2014).
- 4) Tian, Y. and Zhu, Y. : Better Computer Go Player with Neural Network and Long-term Prediction, arXiv:1511.06210 (2015).
- 5) 伊藤毅志, 村松正和 : UEC 杯コンピュータ囲碁大会と電聖戦 2014, 人工知能学会誌, 特集「コンピュータ囲碁」, Vol.29, No.4, pp.311-315 (2014).

(2016年2月8日受付)

■ 伊藤 毅志 (正会員) ito@cs.uec.ac.jp

1994年名古屋大学大学院工学研究科情報工学専攻博士課程修了。工学博士。同年より電気通信大学情報工学科助手。2007年より同助教。電気通信大学エンターテイメントと認知科学研究ステーション代表。UEC 杯コンピュータ囲碁大会実行委員長。ゲームを題材にした人間の思考過程、熟達化の過程に興味を持つ。著書に『先を読む頭脳』(新潮社, 共著)ほか。コンピュータ囲碁フォーラム理事。

■ 村松 正和 (正会員) MasakazuMuramatsu@uec.ac.jp

1994年総合研究大学院大学数物科学研究科統計科学専攻博士課程修了。博士(学術)。同年上智大学理工学部機械工学科助手。2000年より電気通信大学講師, 2002年助教授, 2008年教授。専門は最適化とゲーム情報学。特にコンピュータ囲碁に興味を持ち、伊藤と協力してUEC 杯コンピュータ囲碁大会を開催してきた。コンピュータ囲碁フォーラム副会長。

