

非線形評価関数の改良により 臨んだ第3回将棋電王戦

竹内 章



電王戦出場への思い

1年前に行われた第2回電王戦第1局で、私の開発したコンピュータ将棋ソフト「習甦^{しゅうそ}」は阿部光瑠四段に完敗した。対局直後は、頭の中が真っ白だった。ただ、自分自身に言い聞かせるように思った一言だけは憶えている。「この負けがあったからこそ、今の自分があると言えるようにしたい」。

そのときの無念さは、単に勝負に負けたということではなく、名局を残せなかったという思いからきている。昨年は、コンピュータ将棋の進歩に大きな功績を残し、また実力がありながら不運にも出場権を獲得できなかった「Bonanza」や「激指」を差し置いて、幸運にも出場することができた。コンピュータ将棋開発者にとって貴重な一局であるにもかかわらず、0手で馬を作らせたり、投了近くで無駄な歩の成り捨てをしてしまったりなど、悔いの残る手順ばかりが目立ってしまい、習甦の良さを見せることができなかった。それは、対戦した阿部光瑠四段の戦略が優れていて、習甦の強さを封じ込めたことも理由の1つであったと思う。電王戦は将棋に関心が低い多くの方も注目しており、習甦の強さを分かってもらうことによって、コンピュータ将棋に対する戦略を示した歴史的この一局の価値も高められるのではないかという思いもある。

本稿では、第3回電王戦に向けて行った習甦の改良と、対局中の習甦の思考方法や思考ログの説明を加えながら対局について振り返る。

評価関数の改良

図-1に、習甦の評価関数の概念図¹⁾を示す。3層

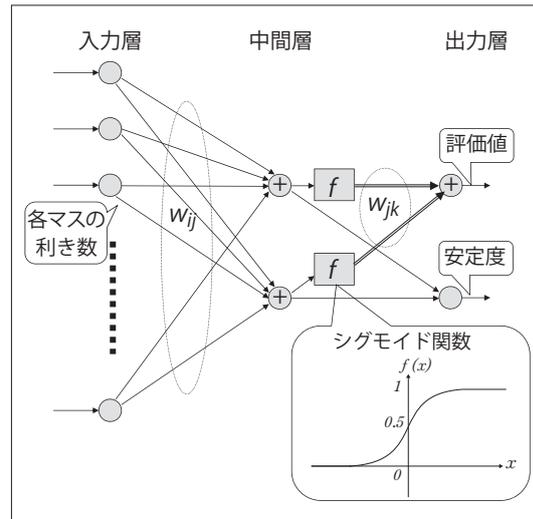


図-1
習甦の評価関数概念図

ニューラルネットワークに似た構造となっており、入力層の情報から盤面全体の各マスに対する利きを評価する。これらの利き評価を足し合わせた値の関数として各駒の価値を算出する。2つの中間層の出力は自玉および相手玉の安全度を意味し、安全度が増すほど自駒の価値が高くなるが、安全にし過ぎても効果は飽和し、危険になると価値がなくなっていく。このような特性はシグモイド関数を用いることで上手く表現できる。利き評価の関数という考え方によって、局面評価における非線形な特性だけでなく、大局観の実現も試みている。

第2回電王戦後の改良として、各駒の価値について、玉との位置関係のみであったものを、あらゆる2つの駒の位置関係に拡張した。評価関数の構造から、自駒と相手の駒との位置関係は、どちらから見たかで2つの価値を持たせるようにした。こうすることによって、2つの中間層を持つ非線形な評価関数をベースとしながら、駒の位置関係という現在主流といえる評価項目を取り入れることにした。

課題となったのは、この評価関数を機械学習させ

第3回将棋電王戦を振り返って

種	9	8	7	6	5	4	3	2	1	
▽	皇	桂		金		金	銀	桂	皇	一
		歩		銀			王	角		二
	歩		歩	歩	歩	歩		歩	歩	三
		歩					歩			四
				歩						五
			歩							六
	歩	歩		歩		歩	歩	歩	歩	七
		角		飛		玉				八
	香	桂	銀	金		金	銀	桂	香	九 ▲菅井

図-2 ▲5五歩まで

種	9	8	7	6	5	4	3	2	1	
▽	皇	桂				金		桂	皇	一
		歩				銀	王			二
	歩			銀	歩	金	角	歩		三
		歩	歩	歩		歩	歩		歩	四
					歩					五
			歩		銀	歩			歩	六
	歩	歩	角	歩			歩	歩		七
			飛				銀	玉		八 ▲菅井
	香	桂		金		金		桂	香	九

図-3 ▲7八飛まで

る際、非線形特有の難しさであった。評価関数の学習において、初期値依存性が強いことは報告されている²⁾。非線形な評価関数においては、それが顕著になるばかりでなく、局所解が複雑に存在し、安定性を維持するのが難しい。学習パラメータの少しの変化で、思わぬ方向へ収束解が向かうこともある。

第3回電王戦への出場をかけた電王トーナメントに向けて、徐々に学習を安定させていった。しかしながら、コンピュータ将棋のトップクラスの層は年々厚くなっており、厳しい対局が続いた。結果的には、前述した電王戦にかける強い思いが幸運を呼んだのか、前回に引き続き最後の出場権を獲得することができた。

第3回電王戦第1局

対局相手は、菅井竜也五段と決まった。菅井五段は、7割を超える勝率で、通算成績において勝率一位を争うほどの安定した好成績を残している新鋭プロ棋士で、対局できるのは本当に光栄なことだと思った。

苦慮した序盤の駒組み

菅井五段は、振飛車党で有名な棋士であり、事前の宣言通り振り飛車を選ばれた。実戦でも、中飛車の最新流行形を得意とする棋士である。図-2が序盤における居飛車側の作戦の岐路となった局面である。

序盤における指し手は、以下に説明するように決定される。まず、各スレッドは、登録された棋譜からランダムに指し手を選択する。選んだ指し手に引

き続き棋譜データベースから指し手を選択して局面を進め、指し手がなくなるとその局面を事前に読ませて記録された評価値を返す。第3回電王戦の統一パソコンは6コアのCPUが搭載されており、ハイパースレッディングをオフとしたため、スレッドは6つであった。図-2の局面では、4スレッドが△5二金右を、2スレッドが△4二銀を選択した。6つの評価値のうち最も高い評価値を返したスレッドが選んだ指し手は△4二銀であった。この△4二銀は最近指されない指し手で、古い棋譜から選んでしまったようである。

その後の局面での読み筋を見ていると、駒組みに苦慮していることがうかがえた。図-3の一手前の局面では、上がった銀を元のマスに戻して穴熊に組み替える順まで読んでいた。しかし、後手番で2手損したのでは過大評価しがちの穴熊に組めたとしても評価値は良くなってはいなかった。

明暗を分けた長考

図-3の局面で、ずっと△2二玉を最善手と読んでいたが、△7二飛に変わった後しばらくして、その手が指された。局後の菅井五段のコメントでは、△7二飛は練習対局では見なかった指し手だったそうである。

局後聞いた話では、持ち時間5時間では一日一局しか指せないため、局数を多くこなすために、持ち時間を減らした条件で練習対局していたそうである。習熟では残り時間に比例した基本時間を設定しているため、持ち時間が短ければ△7二飛ではなく、△2二玉を指した可能性が高い。

9	8	7	6	5	4	3	2	1	
▲	桂					▲	桂	▲	菅井
▽		飛			銀	金	王		二
	歩		銀	歩	金	角	歩		三
		歩	歩		歩	歩		歩	四
				歩					五
			歩	歩	銀	歩		歩	六
歩	歩	角				歩	歩		七
		飛	金		銀	玉			八
香	桂				金		桂	香	九

図-4 ▲3二金まで

図-4は、昼食休憩の局面である。ここですでに、菅井五段は予想外の長考に入っていた。このときの習甦は、▲8八角△5四歩▲同歩△8五歩のような展開を予想していた。この手順の意味は分からないが、評価値 -78 が示すように若干苦戦を意識していた。

休憩明けさらに長考した末に菅井五段が指した手は、果たして▲6八角であった。ここで習甦は、△5四歩▲同歩に続いて、△4五歩の仕掛けが成立すると読み始めた。この手以降、評価値は習甦有利に変化した。

メモリの記憶が生んだ幸運

相手の手番では、相手の指し手を予測してその手が指された局面を読むことが一般的である。習甦は、相手の立場で考え、読んだ局面の評価値や最善手をメモリに記憶していき、自分の手番になったときの読みを効率化するために利用する。電王戦統一パソコンでは、実装メモリは当初 16GByte とされていたが、64GByte に変更された。電王戦のような長い持ち時間の将棋はめったに指せないが、予想以上に効果があるようだった。

図-5の▲4八飛が指される前の局面で、その手に対する応手として、ずっと△5五歩を最善と読んでいた。△5五歩の場合の評価値は84で、途中△4六歩に読み替えて評価値が90と、わずかな差で最善応手が変わっていた。

しばらくして▲4八飛が指され、読みを進めるうちに評価値は90から81、77とわずかではあるが下がっていき、△5五歩の変化の評価値よりも

9	8	7	6	5	4	3	2	1	
▲	桂						▲	桂	菅井
▽		飛				金	王		二
	歩				銀	金	角	歩	三
		歩	歩				歩		四
						歩			五
			歩	歩				歩	六
歩	歩	角	金			歩	歩		七
					飛	銀	玉		八
香	桂				金		桂	香	九

図-5 ▲4八飛まで

低くなっていた。読みが深くなるにつれて評価値が下がっていくときは、悪い結果が待っていることが多い。しかも△4六歩の後の読み筋は▲4六同飛に△7八銀と打ち、桂を取りに行く手順である。駒得（桂）したものの駒の働き（銀）が悪くなり形勢を損ねる、一昔前のコンピュータ将棋が指す悪手のパターンにも見える。

対局中、この△4六歩は嫌な感じで見ていたが、菅井五段は意表をつかれたようである。30分近くの考慮の末、▲6五歩とさばきに出た。ここで、習甦の評価値が初めて100を超えた。

▲4六同飛の変化で評価値が下がっていたことは気になっていて、後日、菅井五段に聞いてみた。明解な返答があり、その後も検討した結果△4六歩は最善手だと結論づけたそうである。ただ、▲4八飛の局面をポンと習甦に読ませてみると、△4六歩は最善手とは読まないようである。

原因は、前述した相手の手番で読んだ局面の評価値や最善手の記憶のほかに、読みの結果良い指し手となった履歴から指し手の読む順序を決定する「ヒストリーヒューリスティック」と呼ばれる方法を採用しており、さらにその順序により読む深さを決めていたこともある。大勝負の実戦の流れにおいて、思考内容がメモリに記憶されたことにより、微妙な手のゆらぎが習甦に幸運をもたらした。

習甦の幸運は続く。図-6の局面で、△5四銀を読んでいたが、評価値は徐々に下がっていた。130まで下がった後、わずか5点の差で最善手が△4二飛に変わった。結果的にこの△4二飛も意表の好手となり、次の▲5七金を見て評価値は200を超えた。

第3回将棋電王戦を振り返って



図-6 ▲3六飛まで



図-7 △6七歩成まで

天来の投了図

96手目△2七銀の局面で習甦は必至や詰みまでの変化を読み切り、思考を停止した。私は、第2回電王戦に続き祖父の形見の和服を着て対局に臨んだが、その祖父が他界したのが96歳であり、何か縁を感じる。そして、メモリに記憶されていた最終手△6七歩成は、まさに天来の鮮やかな決め手になった(図-7)。

第3回電王戦を終えて

対局直後、将棋の内容についての満足感、幸運ながらも勝てた達成感があったものの、何か複雑な気持ちだった。昨年の力を出し切れなかった対局後の無念さを思い出してもいた。プロ棋士のそれは、趣味で開発してきた自分とは比較にならないだろう。

しばらく躊躇していたが、菅井五段に声をかけた。「今後も習甦を使っていただけるなら、最新版の習甦と指していただきたい」という内容だった。阿部四段との交流がなく、またこの対局に勝っていなければ、自分からこういう発言はできなかったと思う。菅井五段の表情から敗戦の色は消え、その後、さまざまな話をさせていただくこととなった。この一局に勝つことだけを考えるなら、方法はあったのだろうと思わせる話も聞いた。しかし、菅井五段の目は将来に向けられていた。

第2局の検討室で聞いた印象に残った言葉がある。「(コンピュータの最善手を見て)こんな手は読めない。しかし、こういう手を何度も見ていると指

せるようになるかもしれない」。同様に、菅井五段の「10年後は人間の方が強い」という発言は、コンピュータ将棋とともに強くなれるという自信の表れだと思う。

第3回電王戦では、視聴者の投票により MVP に選んでいただいた。また、第1局は名局だとも言っていた。観ていて面白かったという点では、ほかの対局のどれもが素晴らしい熱戦だった。敗着不明なことも名局の要素なのだろうか。いずれにしても、名局は1人では作れない。宣言通り振り飛車で真つ向勝負に出て、明確な疑問手を指さなかった菅井五段とともに作り上げた名局だと思う。また、第2回電王戦の敗戦のあと阿部四段からいただいたアドバイスによって、評価関数の大きな改良に早期に踏みきれたことを考えると、阿部四段のお蔭でもあると思う。

電王戦は私の人生の中でも印象に残る、まさに歴史的イベントとなった。電王戦の開催および報道にご尽力いただいた皆様に、改めてお礼を申し上げたい。

参考文献

- 1) 竹内 章：コンピュータ将棋における大局観の実現を目指して、人工知能学会誌, Vol.27, No.4, pp.443-448 (2012).
 - 2) 保木邦仁：第23回世界コンピュータ将棋選手権自戦記, 情報処理, Vol.54, No.9, pp.929-932 (Sep. 2013).
- (2014年6月2日受付)

竹内 章 | akir.takeuch@gmail.com

1990年九州大学工学部電子工学科卒業。1992年同大学院工学研究科電子工学専攻修士課程修了。趣味としてコンピュータ将棋の開発に従事。電子情報通信学会、IEEE各会員。